



Rapid Feedback for Family Planning Improvement

22  
J. Buehler  
P/C  
169-81

# DEMOGRAPHIC TECHNIQUES OF FERTILITY ANALYSIS

*Donald J. Bogue*

*Director, Community and Family Study Center*

*published by*

COMMUNITY AND FAMILY STUDY CENTER

UNIVERSITY OF CHICAGO

1971

FAMILY PLANNING RESEARCH AND EVALUATION MANUAL  
NUMBER

2

12349  
CPHE-CLIC  
R.F.F.P.I. FAMILY PLANNING EVALUATION MANUALS

Number

1. FAMILY PLANNING IMPROVEMENT THROUGH EVALUATION: A MANUAL OF BASIC PRINCIPLES. Donald J. Bogue (editor)
2. DEMOGRAPHIC TECHNIQUES OF FERTILITY ANALYSIS. Donald J. Bogue
3. A MODEL INTERVIEW FOR FERTILITY RESEARCH AND FAMILY PLANNING EVALUATION. Donald J. Bogue
4. TECHNIQUES OF PREGNANCY HISTORY ANALYSIS. Donald J. Bogue and Elizabeth J. Bogue
5. THE CONTRACEPTIVE HISTORY TECHNIQUE FOR MEASURING CONTRACEPTIVE USE EFFECTIVENESS AND IMPACT OF FAMILY PLANNING UPON BIRTH RATES. Donald J. Bogue assisted by James Nelson and Jeanne Cairns
6. A TECHNIQUE FOR PROJECTION OF FAMILY PLANNING TARGETS AND QUOTAS REQUIRED TO ATTAIN DEMOGRAPHIC OBJECTIVES. Donald J. Bogue with computer programs prepared by Scott Edmonds and Mohammed Rafiq
7. MINI-TAB: A PACKAGED CROSS-TABULATION PROGRAM FOR PROCESSING SURVEY DATA ON SMALL COMPUTERS. Henry D. Elkins
8. A SYSTEM FOR MEASURING CONTRACEPTIVE USE-EFFECTIVENESS THROUGH FOLLOW-UP SURVEYS OF ADOPTERS: A SET OF COMPUTER PROGRAMS. Christopher
9. AM FOR AND OTHER EVALUATION.

Other n  
write to  
Univers

Co  
Permis  
produc  
or all  
author

les available,  
y Study Center,  
Illinois 60637.

y Center  
ization to re-  
ublish parts  
ent from the

**SOCHARA**  
Community Health  
Library and Information Centre (CLIC)  
Centre for Public Health and Equity  
No. 27, 1st Floor, 6th Cross, 1st Main,  
1st Block, Koramangala, Bengaluru - 34  
Tel : 080 - 41280009  
email : clic@sochara.org / cphe@sochara.org  
www.sochara.org

## PREFACE

One of the unmet needs experienced by researchers who are engaged in fertility and family planning research is the availability of a single systematic treatise on the demographic methods of fertility measurement. Few evaluation problems occupy the attention of family planning researchers more than the question of how much, if any, birth rates have declined since major family planning programs were introduced. It is the purpose of this manual to place at the disposal of these researchers techniques with which they can exhaust the information contained in census and survey data.

The present manual is a draft version of Chapter 18 in a book, Techniques of Demographic Research, being written in collaboration with Professor Evelyn M. Kitagawa, Associate Director of the Population Research Center of the University of Chicago. The chapter is now undergoing joint review, revision and expansion by the authors; in its final form it will be more detailed and complete than this preliminary version.

As for other publications in this series, this manual was released for use as training material in workshops on family planning evaluation, being held throughout the world under a contract with the U.S. Agency for International Development.

Because a methodological work on this subject has not yet appeared in print, the authors would be most grateful to receive letters from their colleagues containing suggestions both for improving the presentation and for including additional topics.

June 1, 1971

Donald J. Bogue

## ACKNOWLEDGEMENTS

The first draft of these materials was written in 1954. Since then many research assistants, students and other interested persons have participated in research and experimentation to expand and refine the methodology. Their contributions, both in the form of research assistance and in the roles of professional readers and critics is gratefully acknowledged. Three persons have made especially great contributions to this undertaking: Lee-Jay Cho, James Palmore, and Elizabeth J. Bogue. Dr. Wilson H. Grabill, Chief of Fertility Statistics, U.S. Bureau of the Census, Dr. Christopher Tietze, Deputy Director of the Biomedical Division, the Population Council, and Dr. Roy Treadway, Population Institute of Hacettepe University, Ankara, Turkey, read the manuscript and made invaluable suggestions and corrected errors. Their help was particularly welcome because all are experts at extracting fertility estimates indirectly from census or other special sources--a task especially important in family planning evaluation.

Many thanks are due to Mary Teahan, Diane Buczek, Carmen Fernandez, Norma Dutter, Nancy Caparros, and Joyce Bacon for typing and assembling the manuscript.

The research that underlies the preparation of this manuscript was performed under grants for work in basic demography methodology received from the Rockefeller Foundation between 1954 and 1964, and from the Ford Foundation for research on family planning and evaluation from 1965 to the present. Preparation of copy for printing was financed by a contract with the U.S. Agency for International Development.

D.J.B.

## CONTENTS

	Page
18-1. Introduction.....	1
a. Distinction between "fertility" and "reproduction" and between "birth rate" and "fertility rate" .....	1
b. Bodies of data available for fertility analysis.....	2
18-2 Vital Statistics of Birth.....	2
18-3 Basic Measures of Fertility.....	4
a. Crude birth rate .....	4
b. The general fertility rate (GFR) .....	5
c. Age-specific fertility rates (ASFR).....	8
d. Total fertility rate (TFR) .....	10
e. Age cumulative fertility rate (ACFR).....	11
f. Percent distribution of total fertility.....	11
g. Single years of age versus 5-year age groupings.....	11
h. Fertility rates specific for other traits .....	14
18-4 Interrelationships Among the Basic Fertility Measures ...	14
18-5. Cross-Sectional General Fertility Rates Standardized for Age and Other Characteristic of Parents .....	16
18-6 The Measurement and Explanation of Fertility Change ....	18
a. Cross-sectional analysis of fertility change .....	20
18-7. Longitudinal (Real Cohort) Measures of Fertility.....	25
a. The longitudinal fertility cycle.....	25
b. Real current fertility, cumulative fertility, and completed fertility .....	26
c. The need for a longitudinal analysis of fertility .....	26
d. Definition of real cohorts .....	28
e. Basic cohort fertility rates.....	28
f. Estimation of births and woman-years for longitudi- nal rates from conventional vital statistics and cen- sus tabulations.....	29
g. Fiscal real cohorts .....	32
h. Age-groupings of real-cohort data .....	34
i. Example of construction of a real cohort.....	34
j. Cohort analysis of fertility change.....	35
k. Children ever born data and fertility change.....	39
18-8 Central Rates Versus Birth Probabilities .....	41

# CONTENTS

	Page
18-9 Nuptial (Conjugal) and Illegitimacy Fertility Rates .....	42
18-10 Calculation of CBR and GFR from Age Distributions.....	45
18-11 Child-Woman Ratios and Their Relationship to the Basic Fertility Rates.....	56
18-12 Estimation of Basic Fertility Measures from Child- ren Ever Born Data: Cumulative Longitudinal Fer- tility .....	57
a. Children ever born rates.....	58
18-13 Calculation of Current Birth Rates from "Own Children" Tabulations of Census Data .....	65
a. Relationship between age and time.....	66
b. Coding procedure .....	67
c. Tabulation procedure .....	69
d. Adjustment to correct for deficiencies .....	70
e. Conversion of ratios of children to women into esti- mates of ASFR.....	72
18-14 The Measurement and Explanation of Differential Fertility.....	76
18-15 Proportional Fertility or Birth Ratios .....	80
18-16 Birth Order Proportions and Rates .....	81
18-17 Estimation of Basic Fertility Measures by Regression Analysis.....	84
a. Estimation of basic fertility measures by regression analysis: System R.....	85
b. Estimation of basic fertility measures by regression analysis: System O.....	98
18-18 Deficiencies in Vital Statistics and Census Data for Measuring Fertility and Techniques for Computing Correction Factors .....	101
a. Underregistration of births.....	101
b. Underenumeration of children.....	110
c. Underreporting of children ever born.....	114

## Chapter 18

### FERTILITY ANALYSIS

18-1. Introduction. As it is used by demographers, the term "fertility" refers to childbearing. It is to be contrasted with "fecundity," which is the term used to refer to the capacity to bear children. Thus, a fertility rate refers to the relative frequency with which births actually occur within a population. Because it is a major expansionary force in population dynamics and a major counteracting force to population attrition from mortality, fertility is a leading branch of study in the field of demography. The forces causing fertility among a population are more complex than those causing mortality; hence, we need a more elaborate methodology for studying fertility.

Demographers have developed a variety of methods for measuring fertility. All of these measures are interrelated. Unfortunately, the nature of these interrelations has not been fully explored. The present chapter undertakes six tasks. (a) It discusses the sources of fertility data and the limitations of these data. (b) It describes each of the numerous fertility measures that have been developed and illustrates how each is computed. (c) It compares these fertility measures with one another and attempts to define clearly the differences among them. (d) It attempts to simplify and systematize the entire field of fertility analysis with the hope of increasing the comparability of research results among places and over time. (e) It illustrates with examples the major applications of the fertility measures to the study of inter-group and inter-area differences in fertility. (f) It presents both the "real cohort" and "hypothetical cohort" perspectives of fertility analysis and also attempts to reconcile and integrate them.

18-1a. Distinction Between "Fertility" and "Reproduction" and

Between "Birth Rate" and "Fertility Rate." In studying fertility, a distinction must be carefully maintained between "fertility" and "reproduction." Whereas fertility analysis deals only with the childbearing activity of a population, reproduction deals with the ability of a population to replace itself and to grow. The study of reproduction considers the processes of birth and death together and tries to determine the extent to which the rate of childbearing is sufficiently great to counterbalance the force of mortality and permit population growth. In other words, the study of reproduction is the study of the net balance between fertility and mortality in a population. A later chapter deals with reproduction.

In this manual, a distinction is made between "birth rates" and "fertility rates." A birth rate refers to the rate of incidence of births among a general population, while a fertility rate refers to the rate or incidence of births in a female population. Wherever the female population, or some part of it, is used as the denominator of a ratio in which a group of births or children constitutes the numerator, the resulting measure is termed a fertility rate. It has become popular to use the word "natality" as a synonym for "fertility." That practice has not been followed here because it would require a renaming of several measures that traditionally are labeled "fertility" rates.

18-1b. Bodies of Data Available for Fertility Analysis. There are three major sources of fertility data: (a) vital statistics of birth obtained by the registration of each birth event as it occurs, (b) periodic census enumerations, and (c) special fertility surveys. Together these three sources make available five different bodies of data from which fertility measures may be derived, either as direct measures of the incidence of childbearing or as indirect measures of the comparative level of childbearing:

1. Vital statistics.....Registration or repeated sample surveys
2. Age distributions.....Censuses or sample surveys
3. Children-ever-born.....Censuses or sample surveys
4. Own children under 10 years of age.....Censuses or sample surveys
5. Birth rosters.....Sample surveys (potentially a census item).

The first step in learning fertility analysis is becoming familiar with the measurement of fertility using vital statistics. With these data we can establish all of the essential concepts and measures of a basic system. Following this we will learn how to use other sources of data to calculate or estimate these same measures or variants of them.

18-2. Vital Statistics of Birth. Chapter 6 has already described how vital statistics of birth are collected and tabulated. Nations of

Europe and North America almost all have series of birth statistics that range in quality from good to excellent, and several countries possess series of birth statistics that extend backward over more than a century. Good birth statistics also are now available for Japan, Argentina, Israel, Taiwan, and several other nations. The remaining developing nations of Asia, Latin America, and Africa have registration systems that vary from moderately good to extremely poor. Thanks to programs for registration improvement, the quality of registration data is rising rapidly throughout the world, and several nations that just a few years ago had only very inadequate birth data now produce birth data that are usable.

The first nationwide efforts to collect birth statistics in the United States were made in the censuses of 1850, 1860, and 1870, where children under one year of age were taken as an index of births. In the censuses of 1880, 1890, and 1900, these data were corrected by adding an estimate of deaths of persons born in the preceding year to get an estimate of births during the year preceding the census. Due to underenumeration and age misstatement, the results were not satisfactory.

Prior to 1915, the only usable birth statistics for the United States were tabulations made by a few individual states, some of which had been registering births for many years. In 1915 a Birth Registration Area was set up. It consisted of ten states and the District of Columbia. Not until 1933 did the last state meet the stringent requirements (at least 90 percent completeness) for admission to the Birth Registration Area. After 1933, there was gradual improvement, due primarily to an increase in the proportion of births that occurred in hospitals. Since 1960 registration has been nearly 100 percent complete.

In most nations a standard certificate of birth specifies the items of information that must be collected. Samples of such certificates are reproduced in Chapter 6. The following items of information of interest to the demographers may be found on many certificates.

1. Place of birth (place of occurrence)
2. Usual place of residence of mother or parents
3. Sex of child
4. Date of birth
5. Multiple or single birth
6. Ethnicity or race of parents
7. Age of mother
8. Birthplace of mother
9. Birth order (number of children previously born to mother)
10. Religion of parents
11. Age of father
12. Birthplace of father

13. Occupation of father
14. Industry of father
15. Attendant at birth--physician or midwife
16. Occurrence of birth in hospital or outside
17. Legitimacy or illegitimacy of birth
18. Length of pregnancy
19. Birth weight
20. Date of parents' marriage or duration of marriage

As pointed out in Chapter 6, it is highly desirable that the categories of personal characteristics of the infant and his parents (and the definitions used in classification) be identical to those used for collecting census data in order that the fertility rates may have the greatest possible validity.

Vital statistics of birth may be deficient in four ways: incomplete registration, late registration, errors of residential allocation, and misreporting age of mother or other facts on the birth certificate. Incomplete registration is by far the major shortcoming of most birth data derived from vital statistics. Chapter 7 has described general procedures for measuring the extent of this deficiency. A later section in this chapter considers in more detail each of the four deficiencies. Techniques for measuring the extent of underregistration and for arriving at correction factors are described and illustrated.

18-3. Basic Measures of Fertility. All fertility rates consist of a measure of childbearing expressed as a ratio of the number of births that occur in a population to the size of the population that bears the children. Demographers have devised five such rates. These rates are highly correlated with each other, and together they may be viewed as a system of fertility analysis. Several additional measures will be presented, but all will turn out to be elaborations or refinements of one or another of these five:

Crude birth rate (CBR)

General fertility rate (GFR)

Age-specific fertility rates (ASFR)

Total fertility rate (TFR)

Age-cumulative fertility rate (ACFR)

18-3a. Crude Birth Rate. This is the number of births per 1000 midyear total population:

$$CBR = \frac{B}{P} \times 1000 \quad (1)$$

where B is the number of births that occur during a calendar year, and  
P is the midyear total population.

This measure was described and illustrated in Chapter 2; hence, it

needs no further elaboration here. It is a crude measure of child-bearing because the denominator contains a large population not exposed to childbearing: males, children, and elderly persons. Despite this shortcoming, the measure is a highly useful one. We may think of this measure in the following way.

$$\text{CBR} = \frac{B}{E+N} \times 1000 \quad (1a)$$

where E is the "exposed" population, and  
N is the "unexposed" population.

If the quantity N were always to stand in some fixed ratio to E, then changes in CBR would signal changes in fertility very precisely. Because the ratio of N to E does tend to change rather slowly, CBR can prove to be quite useful in measuring short-run changes in fertility within a particular population. It is somewhat less useful in comparing the fertility of two populations at a given moment. Even here, however, it is useful as a general indicator of fertility levels; except in extreme situations the ratio of N to E has only a moderate variation from one population to another. By shifting from the crude rate to one of the more refined measures, demographers often are accomplishing much less refinement than many may think.

A major weakness of this measure is that it is not very sensitive to small fertility changes; in fact, it tends to minimize them. If the birth rate rises, there is an increase in the proportion of children in the population. This swells the size of the denominator and tends to mask the fertility rise. The same phenomenon, with reverse distortion, tends to understate the extent of a genuine fall in fertility. For the same reasons it tends to give a distorted measurement of the timing of fertility change; it will tend to place the peak of a fertility rise somewhat earlier than the more refined measures and will place the trough of a fertility decline somewhat later than indicated by the more refined measures.

18-3b. The General Fertility Rate (GFR). The general fertility rate is defined as the number of births per year per thousand women of childbearing ages.

$$\text{GFR} = \frac{B}{F} \times 1000 \quad (2)$$

where B is the total births that occur during a  
calendar year, and  
F is the total number of females of childbearing  
age (midyear population).

Demographers differ in their definitions of the years to be considered as the childbearing period. Some use the interval 15 to 49. Others use the interval 15 to 44, and still others prefer the ages 20-

44. The United Nations uses the interval 15 to 49 years, presumably to encompass the full range of fertility conditions around the world. The United States National Office of Vital Statistics uses the interval 15 to 44 years because in the United States there is almost no childbearing after age 44. In utilizing published rates one must make certain that the denominators are comparable or are adjusted to be comparable.

So very little childbearing takes place after age 44, even in high fertility populations, that almost no distortion results from attributing all such births to the age group 40-44 years of age. For greatest precision and worldwide comparability, the age interval 15-44 should be adopted as standard, with births under 15 and over 44 allocated to ages 15-19 and 40-44, respectively.

The chief virtue of the GFR measure is that it removes from the denominator most of the population that is not directly exposed to childbearing--males, children, and women who have passed menopause. As a result, it is a much more acceptable measure of fertility levels. It is easy to compute (and the necessary information is always available if there are any vital statistics at all) because it requires no additional information about births. The needed information about women of childbearing age is almost always available from censuses or can be estimated. Although the standardized GFR (see below) is slightly more precise for making interpopulation comparisons, in most cases the gain is modest. The computation of the measure is illustrated in Table 18-1.

The practice of using the female population as the "exposed" population is widely accepted as standard procedure, although males of childbearing age could be used with equal logic. Later in this chapter a separate section is devoted to fertility rates based upon the "exposed" male population.

The chief criticisms of GFR are three: (a) If one is seeking a true probability measure of fertility, it is inadequate because the number of women at risk has been approximated only roughly. Certainly in many nations a minority of the girls 15-19 years of age are exposed to the risk of childbearing. For example, throughout most of Europe far less than one-half of the girls are married by age 19, and of those who marry at 18 or 19, all but a small fraction bear their first child in their 20th year or later. All over the world, only a small fraction of the women aged 45 to 49 are able to bear live children; by this age 80 percent or more have passed menopause or conceive only with difficulty because of lowered fecundity and less frequent sexual intercourse. Including such a large non-exposed population in the denominator damages its probability implications. (b) All unmarried women of all ages are included in the denominator. Were it not

Table 18-1. ILLUSTRATION OF PROCEDURE FOR CALCULATING GFR, ASFR, TFR, CFR, AND MEASURES OF AGE DISTRIBUTION OF CHILDBEARING; FIVE YEAR AGE GROUPS: UNITED STATES POPULATION: 1966

Age of women	Midpoint	Births to mothers of age	Mid-year female population (000)	Age specific fertility rate ASFR	Age cumulative fertility rate ACFR	Percent distribution of fertility		Col (1) times Col (6)
						In age	Cumulative to age	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
15-19 years (a) ....	17.5	629,554	8,806	71.5	357.5	17.5	17.5	306.2
20-24 years ....	22.5	1,297,990	6,981	185.9	1287.0	36.0	53.5	810.0
25-29 years ....	27.5	872,786	5,840	149.4	2034.0	24.2	77.7	665.5
30-34 years ....	32.5	474,542	5,527	85.9	2463.5	13.1	90.8	425.8
35-39 years (b) ....	37.5	252,526	5,987	42.2	2674.5	7.0	97.8	262.5
40-44 years ....	42.5	78,876	6,371	12.4	2736.1	2.2	100.0	93.5
Total, all ages.....	3,606,274	39,512			100.0			2563.5

General fertility rate..... 91.3

Total fertility rate..... 2736.1

Average age of mother at birth of child..... 25.6

(a) Includes 8,128 births to women under 15 years of age.

(b) Includes 4,436 births to women 45 years of age or older.

General fertility rate =  $(3,606,274/39,512,000) 1000 = 91.3$

Age specific fertility rate = entry column (2) divided by entry column (3) times 1000

Age cumulative fertility rate = cumulative addition of entries in column (4) multiplied by 5

Total fertility rate = sum of column (4) multiplied by 5 = entry for age 40-44 in col. (5)

Percent distribution in age = ASFR times 5 divided by total fertility rate

Percent distribution cumulative to age = cumulative addition of entries in column (6)

Average age of woman at birth of child = sum of the products obtained by multiplying by entries in column (6) divided by 1000

Source: U.S. Public Health Service. Vital Statistics of the United States: 1966, Vol. 1. Natality. Data for births derived from Table 1-45; data for female population derived from Table 4-2.

for the fact that in most parts of the world the percentage of women not married at ages 25 and over tends to be small and remains almost constant over time, this would be a serious criticism. Actually, it affects the GFR as a fertility index comparatively little. In some applications it is important not to restrict the analysis to married women, but to express the rates in terms of all women. For a more detailed discussion of this point, see the section on nuptial fertility rates, below. (c) The third major criticism of this measure is that it does not control variations in age composition within the reproductive-age range. This weakness may be corrected by standardization, as described below.

18-3c. Age-Specific Fertility Rates (ASFR). The frequency of childbearing varies markedly from one age group to another within the population. In fact, there is a characteristic age pattern to fertility which is very similar all over the world. This age pattern is best revealed by computing age-specific fertility rates. The age-specific fertility rate is the number of births per year per 1000 women of a specified age. Expressed symbolically, it is defined as

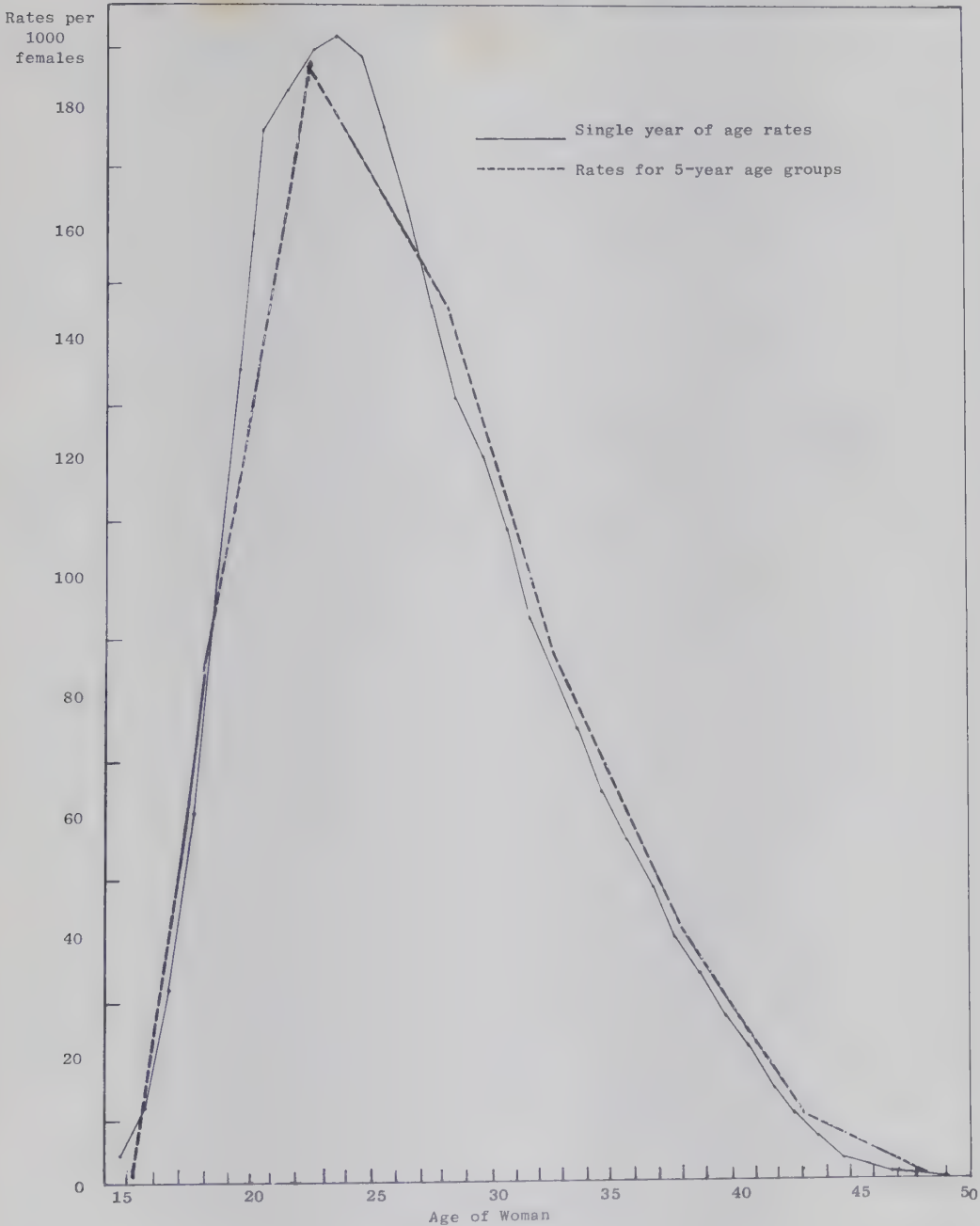
$${}_n f_x = \frac{{}_n B_x}{{}_n F_x} \cdot 1000 \quad (3)$$

where  ${}_n f_x$  is age-specific fertility rate of women aged  $x$  to  $x+n$  years,  
 ${}_n F_x$  is the number of women aged  $x$  to  $x+n$  years at midyear,  
 and  
 ${}_n B_x$  is the number of births to women aged  $x$  to  $x+n$  years during a calendar year.

Age-specific fertility rates may be computed for single years of age or for age intervals. Usually, five-year age intervals beginning with ages of 0 and 5 are employed, as illustrated in Table 18-1. The resulting age-curve of fertility has a very distinctive shape, as illustrated in Figure 18-1. It rises sharply and swiftly from zero at about age 15 to a peak in the early or mid-20's, depending upon the population. Thereafter, it declines somewhat more gradually (but nevertheless rapidly) to reach zero again at about age 50. Populations all over the world manifest this same age-specific fertility pattern. They differ only in the height of the peak rates and the age at which this peak occurs. Given these two pieces of information, it is possible to estimate rather precisely the entire curve.

It is most important that mystic inherent properties not be attributed to the age-specific fertility curve. The shape of this curve is determined by the collective interaction of a definite set of factors, and differences between populations must be understood in terms of differences in these factors:

Figure 18-1. COMPARISON OF AGE-SPECIFIC FERTILITY RATES COMPUTED BY SINGLE YEARS OF AGE AND BY FIVE-YEAR AGE GROUPINGS; UNITED STATES: 1966



- (a) The age-at-marriage distribution.
- (b) The proportion of women of each age who are cohabiting, either within or outside marriage.
- (c) The distribution of sterility and subnormal fecundity by age.
- (d) The age distribution and the birth-order distribution of the use of birth control.

These factors, which are not mutually exclusive, are discussed in more detail below. They are mentioned here only to remind the

reader that age-specific fertility rates are a weighted result of many different interacting variables.

It would be difficult to emphasize too strongly the importance of analyzing fertility in terms of age-specific fertility rates. A schedule of such rates reveals more information about the childbearing activity of the population at a given moment than any of the other measures presented in this chapter. Noting the point-differences in the rates and the relative change in the rate for each group is a detailed and precise way of studying fertility changes. The fact that such an analysis involves a whole schedule of rates and does not yield a single succinct measure should be regarded as an asset rather than as a disadvantage. Noting the ages at which fertility levels are changing is a most informative clue to the forces underlying changes in fertility levels.

18-3d. Total Fertility Rate (TFR). The total fertility rate is a summary over-all measure of fertility obtained by summing the age-specific fertility rates for each year of the childbearing span. It states the number of children 1,000 women would bear during their lifetimes if they were to bear children throughout their lives at the rates specified by the schedule of age-specific fertility rates for a particular year. It assumes that this hypothetical cohort of women suffers no mortality but has the marriage and childbearing patterns of the particular year for which age-specific fertility rates were computed. Because it is a summary measure, the TFR is an alternative to the GFR. Its advantages over the GFR are that it is standardized for age; it contains an equal number of women (1,000) in each age group; hence all TFR's are based upon the same age composition. Another advantage is that it shows the average size of completed family. By dividing by 1,000 (pointing off three decimals), it shows how many children each woman in a population could have borne by the time she passes out of the childbearing ages if a particular schedule of age-specific fertility rates were to determine her childbearing behavior throughout her reproductive life. Table 18-1 illustrates the calculation of the TFR.

$$TFR = n \sum_{15}^{44} \left( \frac{\frac{B_x}{F_x}}{\frac{n}{n}} \right) \cdot 1,000 \quad (4)$$

where data are grouped into age intervals of  $n$  years.

Where rates are for single years, equation (4) becomes:

$$TFR = \sum_{15}^{44} \left( \frac{B_x}{F_x} \right) \cdot 1,000 \quad (4a)$$

In countries where the onset of childbearing is very early or extends very late into the 40's, the range of summation can be altered to fit the circumstances. Some demographers calculate a similar measure, the Gross Reproduction Rate (GRR), which is identical to the TFR except that it refers to female births only. In other words, it is based on ASFR for female births. If we multiply the TFR by the proportion of all births in the year that were female, we will obtain a very close approximation of the GRR. This measure indicates how many daughters 1,000 women would bear during their childbearing years if, throughout their reproductive lives, they were exposed to the age-specific rate of bearing female infants current for a particular interval of time. (The rationale behind the calculation of GRR and its use in demographic analysis will be discussed in the chapter on reproduction.)

18-3e. Age Cumulative Fertility Rate (ACFR). Utilizing the concept of a hypothetical cohort of 1,000 women bearing children throughout their lives at the rate specified by the ASFR for a particular year, we can calculate how many children this cohort of women will have borne by the time its members reach each year of age. This is accomplished simply by cumulating the ASFR for all ages up to the age for which a statistic is desired. (Where the fertility rates are calculated for groups of years, it is necessary to multiply each rate by the number of years in the age-interval before or after cumulating, in order that each year's experience may be given its average weight.) This measure will prove to be highly useful in our later comparisons of hypothetical and real cohorts. Its calculation is illustrated in Table 18-1, column (5).

18-3f. Percent Distribution of Total Fertility. Continuing with the concept of a hypothetical cohort of 1,000 women free from mortality, we can establish three measures which are not rates but which are of value in showing the timing of fertility. First, we can show the percentage of total lifetime fertility that occurs at a particular age or span of ages. This is obtained by multiplying ASFR by the number of years in the age interval and then dividing by TFR. (See column (6) of Table 18-1.) Secondly, we can show the cumulative percentage of fertility--the percentage of total fertility that is completed by the time a particular age is reached. This is accomplished by cumulatively summing the percent distributions for each age interval or by dividing CFR by TFR. Column (7) of Table 18-1 illustrates this measure. Thirdly, we can calculate the average age of a woman at the birth of a child. This is obtained by multiplying the percentage of total births that occur in each age interval by the midpoint of that age interval, then summing across all intervals and dividing the resulting total by 100. Column (8) of Table 18-1 illustrates the procedure.

18-3g. Single Years of Age versus 5-Year Age Groupings. Table

Table 18-2. ILLUSTRATION OF PROCEDURE FOR CALCULATING GFR, ASFR, TFR, CFR AND MEASURES OF AGE DISTRIBUTION OF CHILDBEARING; SINGLE YEARS OF AGE; UNITED STATES POPULATION: 1966

Age	Midpoint	Births to women during this year of age	Midyear female population of this age (000)	Age specific fertility rate ASFR	Age cumulative fertility rate ACFR	Percent distribution of fertility		Col. (0) times col. (5)
						In age	Cumulative to age	
	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
14 years.....	14.5*	8,128	1,810	4.5	4.5	0.2	0.2	2.9
15 years.....	15.5	22,322	1,766	12.6	17.1	0.6	0.8	9.3
16 years.....	16.5	56,634	1,718	33.0	50.1	1.6	2.4	26.4
17 years.....	17.5	107,748	1,735	62.1	112.2	3.0	5.4	52.5
18 years.....	18.5	177,192	1,736	102.1	214.3	4.9	10.3	90.6
19 years.....	19.5	257,530	1,849	139.3	353.6	7.1	17.4	138.4
20 years.....	20.5	241,742	1,380	175.2	528.8	6.7	24.1	137.4
21 years.....	21.5	255,384	1,393	183.3	712.1	7.1	31.2	152.6
22 years.....	22.5	269,672	1,395	193.3	905.4	7.5	38.7	168.8
23 years.....	23.5	283,032	1,485	190.6	1096.0	7.8	46.5	183.3
24 years.....	24.5	248,160	1,326	187.1	1283.1	6.9	53.4	169.0
25 years.....	25.5	220,006	1,228	179.2	1462.3	6.1	59.5	155.6
26 years.....	26.5	193,040	1,177	164.0	1626.3	5.4	64.9	143.1
27 years.....	27.5	170,840	1,170	146.0	1772.3	4.7	69.6	129.2
28 years.....	28.5	155,002	1,155	134.2	1906.5	4.3	73.9	122.6
29 years.....	29.5	133,898	1,119	119.6	2026.1	3.7	77.6	109.2
30 years.....	30.5	121,660	1,106	110.0	2136.1	3.4	81.0	103.7
31 years.....	31.5	104,752	1,110	94.4	2230.5	2.9	83.9	91.4
32 years.....	32.5	91,518	1,082	84.6	2315.1	2.5	86.4	81.2
33 years.....	33.5	82,462	1,105	74.6	2389.7	2.3	88.7	77.0
34 years.....	34.5	74,150	1,134	65.4	2455.1	2.0	90.7	69.0
35 years.....	35.5	67,104	1,156	58.0	2513.1	1.9	92.6	67.4
36 years.....	36.5	58,676	1,178	49.8	2562.9	1.6	94.2	58.4
37 years.....	37.5	48,880	1,198	40.8	2603.7	1.4	95.6	52.5
38 years.....	38.5	43,380	1,223	35.5	2639.2	1.2	96.8	46.2
39 years.....	39.5	34,486	1,245	27.7	2666.9	1.0	97.8	39.5

Table 18-2. ILLUSTRATION OF PROCEDURE FOR CALCULATING GFR, ASFR, TFR, CFR AND MEASURES OF AGE DISTRIBUTION OF CHILDBEARING; SINGLE YEARS OF AGE; UNITED STATES POPULATION: 1966

Age	Midpoint (0)	Births to women during this year of age (1)	Midyear female popula- tion of this age (000) (2)	Age specific fertility rate ASFR (3)	Age cumulative fertility rate ACFR (4)	Percent distri- bution of fertility		Col. (0) times col. (5) (7)
						In age (5)	Cumulative to age (6)	
40 years.....	40.5	27,346	1,250	21.9	2688.8	0.8	98.6	32.4
41 years.....	41.5	19,256	1,265	15.2	2704.0	0.5	99.1	20.8
42 years.....	42.5	13,952	1,278	10.9	2714.9	0.4	99.5	17.0
43 years.....	43.5	8,888	1,290	6.9	2721.8	0.2	99.7	8.7
44 years.....	44.5	4,998	1,295	3.8	2725.6	0.1	99.8	4.4
45 years.....	45.5	2,628	1,251	2.1	2727.7	0.1	99.9	4.6
46 years.....	46.5	1,118	1,225	0.9	2728.6	0.0	99.9	0.0
47 years.....	47.5	448	1,198	0.4	2729.0	0.0	99.9	0.0
48 years.....	48.5	202	1,168	0.2	2729.2	0.0	99.9	0.0
49 years.....	49.5	40	1,140	0.0	2729.2	0.0	99.9	0.0
Total, all ages.....		3,606,274	47,339			99.9		
General fertility rate: Based on females aged 15-44.....				91.3				
Based on females aged 15-49.....				79.2				
Total fertility rate.....					2729.2			
Average age of mother at birth of child.....								25.0
* This figure is for "under 15 years"								

Source: U.S. Public Health Service. Vital Statistics of the United States: 1966. Vol. I, Natality. Table 1-17.

18-1 illustrates all of the measures computed in terms of 5-year age groupings. Table 18-2 illustrates the same measures computed in terms of single years of age. In order to facilitate comparison, data for the same year (1966) are used. Figure 18-1 compares the two distributions graphically. By comparing these tables we may derive the following insights.

- (1) The over-all summaries GFR, TFR, CFR, and average-age-at-birth of a child are affected very little by the use of 5-year groupings rather than the single-year groupings.
- (2) The greatest value of the single-year-of-age statistics is for showing the details of the pattern from age 15 through age 29, where there is very rapid year-to-year change.
- (3) Age heaping and other errors of age reporting may be present. These tend to create irregularities which reduce the meaningfulness of single-year-of-age data. This is especially true for the older ages. These errors may become very serious for low-education populations where the level of accuracy in age reporting is low. Therefore, if single-year-of-age rates are to be meaningfully computed, it may be necessary to smooth or otherwise adjust the age data both for births and for females before calculating the rates.
- (4) Ten-year age intervals are too long to yield precise results.

18-3h. Fertility Rates Specific for Other Traits. Fertility rates may be made specific for any trait for which both the births and the population to which they refer may reliably be classified. For example, it is possible to compute:

- (a) Sex-specific fertility rates (male, female births)
- (b) Ethnic- or race-specific fertility rates
- (c) Order-specific fertility rates
- (d) Duration-of-marriage specific rates
- (e) Marital-status specific fertility rates--for married, single, etc.

Rates that are specific for two or more of these traits can be calculated simultaneously if appropriate cross-tabulations are available both for births and for the female population. Sample surveys that incorporate a pregnancy history offer great opportunities for this.

A difficulty encountered in computing many of these rates is that appropriate denominators are available only for years in which a census is taken; for the intervening years denominators must be estimated. (See chapter on intercensal and postcensal estimates of population.)

18-4. Interrelationships among the Basic Fertility Measures. All of the basic fertility measures purport to measure the forces of fertility upon the total population or some subgroup of it.

1. The GFR and TFR differ only in minor respects:

$$\text{TFR} = \sum_{15}^{44} \frac{B}{F_x}$$

$$\text{GFR} = \frac{B}{{}_3\text{oF}_{15}}$$

For GFR the age groups of birth and females are summed before the division is performed; for TFR the sequence is merely reversed and the division is made before the summation. Under these circumstances, one would expect a very high correlation between GFR and TFR.

2. GFR can be written as  $B/(P-N)$ , where  $P$  is the total population and  $N$  is the population that is not female or not between the ages 15 and 45. Thus, CBR and GFR differ only by the entity  $N$ . Yet demographic analysis teaches us that the ratio  $N/P$  is itself a function of the fertility level because age structure (a function of fertility) is the principal variable determining  $N$ . Thus, we should expect a close correlation between CBR and GFR.
3. If CBR is highly correlated with GFR, it should also be highly correlated with TFR because of the demonstration made in (1) above.
4. Each ASFR is a component of TFR. One should therefore expect high correlations between each ASFR and TFR on the principle of common elements, or part-whole correlation. This correlation should be highest for ages 20-34, where ASFR rates are high, and to be lowest at the extremes of the age span of reproduction.
5. ASFR that are high at one age in a particular population may be expected to be high at all other ages in that population, on the principle that only a lifetime of high fertility can lead to a high level of general fertility. A similar line of reasoning leads us to expect low values of ASFR when GFR is low. We may therefore expect the values of ASFR to be highly correlated with each other as well as with TFR, GFR, and CBR.

The above points constitute an argument that every fertility measure should, by definition, be highly correlated with every other. This is indeed the case. Table 18-3 presents correlation measures between the fertility measures for 50 nations of the world for which the data are of high quality. It is readily apparent that these relationships between GFR and TFR are almost unity, and they are extremely high for all of the measures. Only ASFR for the youngest and oldest ages (where the part-whole relation is most attenuated) have correlations

that are somewhat lower.

Table 18-3. LINEAR CORRELATIONS BETWEEN BASIC FERTILITY MEASURES: 50 NATIONS OF THE WORLD FOR WHICH RELIABLE STATISTICS WERE AVAILABLE FOR THE PERIOD 1955 TO 1960

Fertility measure	Summary measures			Age-specific fertility rates (age)						
	CBR	GFR	TFR	15-19	20-24	25-29	30-34	35-39	40-44	45-49
Summary measures:										
Crude birth rate.....	—	.99	.98	.80	.85	.92	.91	.90	.82	.75
General fertility.....	.99	—	.99	.79	.88	.94	.93	.91	.83	.72
Total fertility.....	.98	.99	—	.78	.88	.96	.94	.91	.83	.71
Age-specific rates:										
15-19 years.....	.80	.79	.78	—	.87	.63	.57	.53	.42	.45
20-24 years.....	.85	.88	.88	.87	—	.81	.68	.64	.53	.48
25-29 years.....	.92	.94	.96	.63	.81	—	.93	.89	.79	.70
30-34 years.....	.91	.93	.94	.57	.68	.93	—	.97	.91	.73
35-39 years.....	.90	.91	.91	.53	.64	.89	.97	—	.95	.75
40-44 years.....	.82	.83	.83	.42	.53	.79	.91	.95	—	.69
45-49 years.....	.75	.72	.71	.45	.48	.70	.73	.75	.69	—

Source: Donald J. Bogue and James A. Palmore, "Some Empirical and Analytic Relations Among Demographic Fertility Measures, with Regression Models for Fertility Estimation," *Demography* Vol. 1, 1964, pp. 319-20.

It is possible to write a linear equation measuring the relationship of any of the basic summary measures to any of the other basic measures. The following are six equations that will be highly useful in translating one form of fertility measurement into other forms:<sup>1</sup>

$$\begin{aligned}
 \text{TFR} &= 30.195 \text{ GFR} + 343.28 \\
 \text{GFR} &= 0.0328 \text{ TFR} - 10.305 \\
 \text{CBR} &= 0.2141 \text{ GFR} + 2.2903 \\
 \text{GFR} &= 4.5952 \text{ CBR} - 8.5945 \\
 \text{CBR} &= 0.0070 \text{ TFR} + 0.2453 \\
 \text{TFR} &= 137.94 \text{ CBR} + 106.16
 \end{aligned}$$

It should be noted that, in theory, the intercepts should all be zero; zero fertility with one measure should imply zero fertility for every other. A study of the intercepts reveals that the ratio of the intercept to the mean value is small and approximates zero.

Table 18-4 has been prepared to assist in making this translation without need for computing; by entering the table with a value of CBR, GFR, or TFR, it is possible to read off the value of all three measures. Regression equations relating these three measures to ASFR and CFR are not presented here because in a later section we present equations which express these relationships even more precisely.

18-5. Cross-Sectional General Fertility Rates Standardized for Age and Other Characteristics of Parents. It was noted above that a weakness of the general fertility rate is that variations of age composition within the reproductive age span are not controlled. This weakness may be remedied by age standardization, using either direct or

Table 18-4. TABLE OF EQUIVALENT VALUES OF BASIC FERTILITY MEASURES BASED ON EMPIRICAL RELATIONSHIPS AMONG THEM

Crude birth rate	Equivalent value			Crude birth rate	Equivalent value		
	GFR (15-49)	GFR (15-44)	TFR		GFR (15-49)	GFR (15-44)	TFR
10.....	37	50	1490	40.....	175	187	5620
11.....	42	54	1620	41.....	180	192	5760
12.....	47	59	1760	42.....	184	197	5900
13.....	51	63	1900	43.....	189	201	6040
14.....	56	68	2040	44.....	195	206	6180
15.....	60	73	2180	45.....	198	210	6310
16.....	65	77	2320	46.....	203	215	6450
17.....	70	82	2450	47.....	207	220	6590
18.....	74	86	2590	48.....	212	224	6730
19.....	79	91	2730	49.....	217	229	6870
20.....	83	95	2860	50.....	221	233	7000
21.....	88	100	3000	51.....	226	238	7140
22.....	92	105	3140	52.....	230	243	7280
23.....	97	109	3280	53.....	235	247	7420
24.....	102	114	3420	54.....	240	252	7550
25.....	106	118	3550	55.....	244	256	7690
26.....	111	123	3690	56.....	249	261	7830
27.....	115	128	3830	57.....	253	266	7970
28.....	120	132	3970	58.....	258	270	8110
29.....	125	137	4110	59.....	263	275	8240
30.....	129	141	4240	60.....	267	279	8380
31.....	134	146	4380	61.....	272	284	8520
32.....	138	151	4520	62.....	276	289	8660
33.....	143	155	4660	63.....	281	293	8800
34.....	148	160	4800	64.....	285	298	8930
35.....	152	164	4930	65.....	290	302	9070
36.....	157	169	5070				
37.....	161	174	5210				
38.....	166	178	5350				
39.....	171	183	5490				

Source: Developed from equations reported in Donald J. Bogue and James A. Palmore, "Some Empirical and Analytic Relations Among Demographic Fertility Measures, with Regression Models for Fertility Estimation," *Demography*, Vol. 1. 1964, p. 219.

indirect standardization as outlined in Chapter 8. If desired, the general fertility rate may be standardized simultaneously for two or more characteristics, such as age and race, age and ethnicity, etc. Table 18-5 illustrates direct age standardization for GFR for three European nations: England, Sweden, and Italy. The estimated age composition of the world female population of reproductive age is used as a standard. It may be seen that there is less variation among the standardized GFR for these countries than among the unstandardized, demonstrating that a substantial part of the differences among their unstandardized GFR are owing to differences in age composition. (As

Table 18-5. ILLUSTRATION OF PROCEDURE FOR DIRECT AGE-STANDARDIZATION OF THE GENERAL FERTILITY RATE: ENGLAND, SWEDEN AND ITALY STANDARDIZED TO AVERAGE AGE COMPOSITION OF WORLD'S POPULATION, 1960

Age	All unmarried women (000) (1)	Fertility rates			Expected births		
		England (2)	Sweden (3)	Italy (4)	England (5)=(X2)	Sweden (6)=(1X3)	Italy (7)=(1X4)
15-19 years.....	195,827	28.3	31.8	21.2	5,542	6,227	4,152
20-24 years.....	174,139	151.2	128.5	103.8	26,330	22,377	18,076
25-29 years.....	153,801	164.5	145.8	150.3	25,300	24,424	23,116
30-34 years.....	139,763	98.3	93.9	111.6	13,739	13,124	15,598
35-39 years.....	125,761	44.6	44.8	56.4	5,609	5,634	7,093
40-44 years.....	111,878	13.5	13.4	27.1	1,510	1,499	3,032
45-49 years.....	98,231	0.8	1.0	2.0	78	98	106
All ages.....	1,000,000	68.9	61.1	71.2	78,108	71,383	71,263

Standardized GFR = (Expected births / standard population) x 1000,  
 England = 78.1 Sweden = 71.3 Italy = 71.3

a country with a long history of low fertility, Sweden has an older age composition; while as a country with very recent fertility decline, Italy has a younger age composition.) The results of this demonstration should warn against interpreting small differences in GFR as genuine differences in fertility until differences in age composition are controlled.

Table 18-6 illustrates indirect standardization. In this instance, the GFR for two cities, New York and Los Angeles, are indirectly multiply-standardized for age and race. Both cities have values of GFR that are only about 75-80 percent as high as the national average. Before standardization, the GFR of Los Angeles was 5.28 points higher than that of New York. After standardization the difference increased to 6.29 points. Thus, none of the differences in fertility level between New York and Los Angeles can be attributed to differences in age and race composition. In fact, compositional differences appear to have concealed a part of the true difference. Both of the above examples also illustrate the principle that very often the differences in fertility that one finds between places in the same region or in the same country are quite small--5 to 10 percent or even less. When attempting to establish the validity of such small differences, it is imperative that one control for the effects of composition.

18-6. The Measurement and Explanation of Fertility Change. All over the world there is great interest in the problem of measuring fertility change. Many countries with high birth rates are anxiously hoping for valid indications that birth rates have started to fall in the wake of national programs of family planning. Many developing countries are studying fluctuations in their birth rates that appear to result from fluctuations in economic conditions, marriage patterns, or

Table 18-6. ILLUSTRATION OF PROCEDURE FOR INDIRECT STANDARDIZATION OF THE GENERAL FERTILITY RATE FOR AGE AND RACE; NEW YORK CITY AND LOS ANGELES, 1960

Age and Race	Standard rates U.S. 1960	Age composition		Expected births	
		New York City	Los Angeles	New York City	Los Angeles
<u>White</u>	(1)	(2)	(3)	(4)	(5)
15-19 years.....	79.8	214,823	63,273	17,143	5,049
20-24 years.....	252.8	213,242	61,791	53,908	15,621
25-29 years.....	194.9	212,769	63,435	41,469	12,364
30-34 years.....	109.6	224,536	72,292	24,609	7,923
35-39 years.....	54.0	235,040	81,510	12,692	4,402
40-44 years.....	15.5	239,099	75,322	3,706	1,168
<u>Nonwhite</u>					
15-19 years.....	162.2	37,282	13,223	6,047	2,145
20-24 years.....	294.2	47,042	16,681	13,840	4,908
25-29 years.....	214.6	50,405	18,174	10,817	3,900
30-34 years.....	135.6	55,649	18,396	7,546	2,494
35-39 years.....	74.2	54,428	17,814	4,039	1,322
40-44 years.....	23.7	45,732	14,763	1,084	350
Total.....	118.0	1,630,047	516,674	196,900	61,646

Expected births: New York = col. (1) x col. (2); Los Angeles = col. (1) x col. (3)

Expected births: New York = Sum col. (4)

Los Angeles = Sum col. (5)

Indirectly standardized birth rate: Adjustment factor

Registered births: New York City = 147,530 Registered/expected = .7495  
Los Angeles = 49,490 Registered/expected = .8028

Standardized GFR: New York City = .7495 x 118.0 = 88.44 Difference = 6.29  
Los Angeles = .8028 x 118.0 = 94.73

Unstandardized GFR:

New York City =  $\frac{147,530}{1,630,047} = 90.50$   
Los Angeles =  $\frac{49,490}{516,674} = 95.78$  Difference = 5.28

timing of childbearing, and trying to disentangle these fluctuations from more fundamental changes in fertility levels. The problem is not an easy one, and demographers disagree concerning techniques and interpretations. Those who are most committed to the study of real cohort fertility place very little confidence in cross-sectional measures. When carried to its logical extreme this position claims that one cannot make inferences about fertility until each cohort has completed childbearing--which precludes even trying to make an interpretation of what is happening at a given time. The opposite error, that of assuming that all fertility change will be spread evenly over all cohorts, is equally untenable. The position developed here is intended to be an intermediate one that will permit inferences to be made about current trends while guarding against serious error.

The methodology for measuring fertility change is very nearly identical for high-fertility and low-fertility populations. There is not one methodology that is appropriate for "developing countries" and

another that is appropriate for "developed countries." The major difference between the two types of population lies in the availability of basic data. Where reliable data are lacking (usually in developing countries) it is necessary to go through a prior estimating procedure to develop the basic measures. Techniques for doing this will be presented below. With this preliminary estimating phase performed, the procedures for measuring change are alike.

18-6a. Cross-Sectional Analysis of Fertility Change. All studies of fertility change should begin with a cross-sectional analysis to determine how much year-to-year change there has been in fertility levels. Once this has been established, a longitudinal analysis may be undertaken to discover how much of this fluctuation may be due to changing childbearing patterns in individual cohorts.

The cross-sectional measurement of fertility change involves the comparison of the basic fertility measures for two or more intervals of time. If the measures are identical for both dates, or are so nearly identical that the differences between them could be attributed to sampling error, then we must conclude that there is no evidence of change. If, however, the measures for the later date are lower than those for the earlier date, we infer that a fertility decline has taken place. Conversely, if the rates at the later date are significantly higher than those at an earlier date, we must infer that a fertility rise has taken place.

The procedure for measuring fertility change in this way is illustrated in Table 18-7. When carried out in its fullest detail, the basic procedure involves the computing or estimating of the values of four fertility measures for each date:

General Fertility Rate (GFR)

Age-Specific Fertility Rates (ASFR)

Total Fertility Rate (TFR)

Age-Standardized General Fertility Rate (GFR-St)

With these values established, two steps are performed to measure change.

- (a) Absolute change--subtract the values for the earlier date from the values for the later date.
- (b) Relative change--express the change in relation to the value at the base (earlier) date. This may be done in one of two ways. Both are acceptable.
  - (1) Percent change--divide the absolute differences, obtained in (a) above, by the value at the earlier date and multiply by 100. This expresses the difference as a percent change.
  - (2) Ratio to base--divide the value of each fertility rate at the later date by the value of the corresponding rate at

the earlier date. This expresses the rate at the second date as a ratio of the rate at the earlier date. A value of 1.00 means fertility decline, and a value above 1.00 means fertility rise. A value of 1.00 means "no change." By subtracting 1.00 from this value and multiplying by 100, we can convert ratios to percent-change measures.

Table 18-7. ILLUSTRATION OF THE "TOTAL" DEMOGRAPHIC PROCEDURE FOR MEASURING FERTILITY CHANGE; DATA FOR THE UNITED STATES: 1960 AND 1965

Age	Age specific fertility rates		Change in rates: 1960 to 1965		
	1965	1960	Absolute	Relative	
				Percent change	Ratio
15-19 years.....	71.2	89.9	-18.7	-20.8	0.792
20-24 years.....	196.8	258.1	-61.3	-23.8	0.762
25-29 years.....	162.5	197.4	-34.9	-17.7	0.823
30-34 years.....	95.0	112.7	-17.7	-15.7	0.843
35-39 years.....	46.4	56.2	- 9.8	-17.4	0.826
40-44 years.....	12.8	15.5	- 2.7	-17.4	0.826
45-49 years.....	0.8	0.9	- 0.1	-11.1	0.889
General fertility rate <sup>a/</sup>	96.6	118.0	-21.4	-18.1	0.819
Total fertility rate..	2928	3654	-726	-19.9	0.801
Age-standardized GFR					
Direct standard.....	94.6	118.0	-23.4	-19.8	0.802
Indirect standard...	94.3	118.0	-23.7	-20.1	0.799

<sup>a/</sup> GFR for U.S. is computed on the basis of women aged 15 to 44; most nations of the world use 15 to 49 or 10 to 49.

The example of Table 18-7 was calculated using data for the United States. It is clear that irrespective of how they are measured fertility rates in this nation declined sharply during the five-year period between 1960 and 1965. The standardized GFR used the 1960 age composition as standard and indicated a decline of 19.8 percent. Almost identical results were indicated by the indirectly standardized GFR and TFR. The relative measures for ASFR reveal that the decline was sharpest among the younger age groups.

When measuring fertility change, the selection of the standard population is of considerable importance. The best standard to use would be the age distribution of the population at the middle of the time span over which fertility change is measured. (This could be an average of composition at beginning and terminal dates.) Barring this, the age composition at the beginning of the interval should be used.

There is a temptation among demographers to overemphasize relative fertility declines and to fail to appreciate the significance of absolute fertility change. The absolute differences measure the absolute impact of fertility change upon population. This is most clearly seen in the instance of the total fertility rate. The absolute decline of 726 per 1,000 women of childbearing age implies a decrease in the average size of completed family of 0.7 children per woman. The absolute change in GFR claims that there were 21 fewer births per 1,000 women of childbearing age in 1965 than in 1960.

In the example given above, it will be noted that the various summary measures do not necessarily give identical results. Under some situations, the differences in results may be much greater than in this example. The question arises: Which of the measures should be accepted as "correct"? It is recommended that they be given priority in the following order:

Standardized GFR (directly standardized)

Total Fertility Rate or GFR (indirectly standardized)

General Fertility Rate (unstandardized)

The TFR is also an age-standardized rate; hence, it is superior to GFR unstandardized. A measure of fertility change should not be regarded as "refined" unless it holds constant the age composition of women within the childbearing ages.

Changes in Crude Birth Rate. The question arises: How much refinement is gained by the procedures outlined above in comparison with simply reporting annual crude birth rates. Using the U.S. example, it may be noted that the crude birth rates for the two dates were as follows:

1960 CBR....	23.7
1965 CBR....	<u>19.4</u>
Change.....	- 4.3

Percent change....	-18.1
Ratio, 1965/1960..	81.9

These results illustrate the principle that for most populations which have not had recent fluctuations either in fertility or in migration over short intervals of time, the relative change in fertility as estimated by the CBR will tend to be nearly the same as the unstandardized GFR. For longer periods of time and where there have been rapid changes either in fertility or migration or both, the CBR can give results quite different from those of the more refined measures. Thus, CBR is least reliable in the very circumstances in which measurement of change is most important.

Standardized Crude Birth Rate. The CBR is such a popular and well-known measure that many demographers would like to express fertility changes in terms of it. CBR has the further advantage of

being the fertility measurement most closely akin to population-growth measures; changes in fertility expressed in terms of CBR can be directly translated into measures of impact upon population growth. In order to achieve these advantages while escaping its weaknesses, it is possible to express fertility changes in terms of CBR standardized for age and sex. The age composition of women of all ages, as of the

Table 18-8. EXAMPLES OF MEASUREMENT OF FERTILITY CHANGE IN DEVELOPING COUNTRIES; PUERTO RICO, WEST MALAYSIA AND KOREA

Fertility measure	Earlier date	Later date	Change	
			Absolute	Percent
PUERTO RICO: 1960-67				
15-19 years.....	101.6	81.7	-19.9	-19.6
20-24 years.....	287.2	208.1	-79.1	-27.5
25-29 years.....	243.2	189.7	-53.5	-22.0
30-34 years.....	157.5	117.8	-39.7	-25.2
35-39 years.....	110.0	79.1	-30.9	-28.1
40-44 years.....	51.1	29.7	-21.4	-41.9
45-49 years.....	9.6	7.5	- 2.1	-21.9
Standardized GFR.....	n.a	n.a	n.a	n.a
General fertility rate....	n.a	n.a	n.a	n.a
Total fertility rate.....	4805	3572	-1233	-25.7
Standardized CBR.....	33.5	25.2	- 8.3	-24.8
Unstandardized CBR.....	33.5	26.9	- 6.6	-19.7
WEST MALAYSIA: 1957-66				
15-19 years.....	173	109	-64	-19
20-24 years.....	285	237	-48	-29
25-29 years.....	276	255	-21	-23
30-34 years.....	254	241	-13	- 5
35-39 years.....	168	163	- 5	- 4
40-44 years.....	57	87	+20	+53
Standardized GFR.....	n.a	n.a	n.a	n.a
General fertility rate....	n.a	n.a	n.a	n.a
Total fertility rate.....	6065	5460	-605	-10
Standardized CBR.....	42.7	37.1	- 5.6	-13
Unstandardized CBR.....	42.9	38.6	- 4.3	-10.0
KOREA: 1957-66				
15-19 years.....	38	12	-26	-68
20-24 years.....	232	206	-26	-11
25-29 years.....	337	310	-27	- 8
30-34 yeras.....	283	252	-31	-11
35-39 years.....	232	163	-69	-30
40-44 years.....	118	85	-33	-28
45-49 years.....	18	15	- 3	-17
Standardized GFR.....	n.a	n.a	n.a	n.a
General fertility rate....	n.a	n.a	n.a	n.a
Total fertility rate.....	6290	5215	-1075	-17
Standardized CBR.....	42.9	35.6	- 7.3	-17
Unstandardized CBR.....	43.8	35.7	- 8.1	-18

Source: "Progress and Problems of Fertility Control Around the World," *Demography*, 5 (2) 1968. Data for Puerto Rico extracted from Jose L. Vasquez, "Fertility Decline in Puerto Rico: Extent and Causes," pp 855-865. Data for Malaysia, from Lee-Jay Cho, James A. Palmore and Lyle Saunders, "Recent Fertility Trends in West Malaysia," pp 732-744. Data for Korea from Lee-Jay Cho and Man Jun Hahm, "Recent Change in Fertility Rates of the Korean Population," pp 690-698.

earlier period, is taken as standard. The age-specific birth rates for the later date are applied to this age composition to get expected births. The sum of the expected births is then divided by the total population of the standard population (males and females of all ages). Alternatively, expected births may be divided by the total standard female population of all ages; the result thus obtained is then divided by the proportion of the total population (both sexes) that was female as of the date to which the standard population refers.

Table 18-8 represents measurement of fertility change in three developing nations: Puerto Rico, Malaysia and Korea. In each case, the authors make use of the model presented above, with the exception of the GFR and the age-standardized GFR.

Often fertility changes occur simultaneously with other changes, and there may be some confusion as to whether or not the changed fertility levels are due to these other changes. One of the most common such changes is the age at marriage. If birth rates begin to fall in a high fertility population, there is likely to be a decrease in child marriage. It then may be unclear whether the lower fertility is due to increased use of contraception or simply to postponement of marriage, with consequent postponement of childbearing. The solution, of course, is to compute the four basic measures as nuptial or conjugal rates, and to measure changes in nuptial fertility rather than in total fertility. (For a discussion of nuptial fertility, see Section 18-9, below.) This should be accompanied by a careful study of the proportions married at each age to determine just how much change in age composition of the exposed population has taken place. Table 18-9 illustrates such computations for the United States. (The rates used computing GFR(N) and TFR(N) are those presented in Table 18-13.)

Table 18-9. CHANGE IN NUPTIAL FERTILITY IN THE UNITED STATES: 1960-65

Age	1960	1965	Difference	
			Absolute	Percent change
15-19 years.....	483.5	452.9	-30.6	-6.3
20-24 years.....	354.4	279.7	-74.7	-21.1
25-29 years.....	222.3	178.9	-43.4	-19.5
30-34 years.....	123.3	101.1	-22.2	-18.0
35-39 years.....	61.7	50.1	-11.6	-18.8
40-44 years.....	17.4	14.1	- 3.3	-19.0
GFR(N).....	156.3	131.1	-25.2	-16.1
GFR(N) stan-				
dardized.....	156.3	126.4	-29.9	-19.1
TFR(N).....	403.0	332.0	-71.0	-17.6

Holding marriage status constant across groups, it appears that very little of the decline in age-standardized GFR can be attributed to changes in marriage patterns. (Compare Tables 18-7 and 18-9.) GFR(N) standardized for age shows only a slightly smaller percentage decline than GFR standardized for age. TFR(N) shows a somewhat smaller relative decline than TFR, primarily because of the small relative change in the fertility of married girls under 20 years of age.

18-7. Longitudinal (Real Cohort) Measures of Fertility. The fertility measures discussed in Section 3 refer to a cross-section of the population during a period of time. They are therefore called period or cross-sectional rates. The technique of building up estimates of the number of children a cohort of women would bear in their lifetime (TFR), or up to a certain age (CFR), if they were to bear children in accordance with a particular schedule of period age-specific fertility rates, is one of constructing hypothetical cohorts. However, we may construct these same five basic measures in terms of real longitudinal cohorts.

Longitudinal fertility analysis keeps a record of (or estimates) the births that occur year after year to groups of women as they pass through the childbearing ages. From these data, collected over the 30 or 35-year age span from the woman's 15th to her 45th or 50th birthday, longitudinal fertility measures are calculated.

Although the longitudinal analysis of fertility and reproduction has its roots early in demographic history, it is only recently that demographers have begun intensively to study fertility from the real cohort, or longitudinal, point of view. An early large-scale empirical study was that of Depoid (1941). Shortly thereafter a number of studies appeared almost simultaneously in several nations. The Basic Readings list the major ones. From this list it can be seen how, within a period of about eight years, longitudinal studies of fertility and reproduction appeared for France, Sweden, Norway, England and Wales, India, and the United States. Whelpton, Woofter, and Ryder have been the leading exponents of this approach in the United States.

The present chapter only introduces this topic; a later chapter is devoted entirely to the longitudinal study of fertility and reproduction.

18-7a. The Longitudinal Fertility Cycle. Assume, as a preliminary definition, that for longitudinal analysis of fertility a real cohort is defined as a group of women born in a given calendar year. The object of longitudinal fertility analysis is to observe how many children this group of women has borne during its lifetime.

From the moment of birth onward, the members of a real cohort are exposed to the likelihood of dying; the probability of death at each age is that specified by a female generational life table for the cohort.

As they approach middle and late adolescence, some of the group begin to marry and then to have children. As the cohort approaches and enters the early twenties, it marries in increasing numbers; by the time it has reached age 30, all but a small fraction have married. Within one or two years after marriage a certain proportion of women have a child, and of these a certain smaller proportion have a second child. Successively smaller proportions will proceed to have their third, fourth, and higher-order children. As the group goes through its 30's, the minority of never-married women is reduced still further by marriage, and the married women continue adding to the cohort's accumulating fertility by adding to their families.

A certain proportion of those who bear no children (though married) are involuntarily infertile, while the remainder have no children because they consciously prevent conception or birth. After one child has been borne, a certain additional proportion of women may begin efforts to prevent further conceptions. This fraction may increase after each additional child has been borne.

Throughout their married lives, the women are subject to loss of their husbands by death or divorce. Of those whose marriages are thus disrupted, a certain proportion remarry after varying lengths of time.

Finally, as the cohort enters its late 30's, its members begin to lose their childbearing potentiality as they enter menopause. By the time age 50 is attained, all but a tiny fraction are no longer fertile, and the childbearing history of the cohort comes to a close. Meanwhile, other real cohorts of women are passing through each phase of the longitudinal fertility cycle to maintain a steady stream of replacements for those passing on to an older age.

18-7b. Real Current Fertility, Cumulative Fertility, and Completed Fertility. The fertility of a real cohort may be expressed in two principal ways: (a) its current fertility, births in a particular year of age; or (b) its cumulative fertility, births up to a specified age since the beginning of childbearing. When all members have passed through the childbearing span (when current fertility has declined to zero), the cumulative fertility becomes completed generational fertility, or the number of children borne by the real cohort in its entire fertile lifetime.

18-7c. The Need for a Longitudinal Analysis of Fertility. Throughout its reproductive life, a real cohort of women bears children under conditions that are not exactly duplicated by any other real cohort.

(1) It tends to marry or to remain single in proportions which may not be exactly duplicated by cohorts that precede or follow it.

(2) Those who marry tend to do so at a distribution of ages that may be different from that of earlier or later cohorts.

(3) Those who marry tend to bear zero, one, two, three, etc. children in proportions that may be different from other cohorts. This

can lead to a different level of completed fertility for the cohort.

(4) The timing of childbearing may differ from cohort to cohort. Some cohorts may wish to start childbearing immediately after marriage and space their children closely together, while other cohorts may prefer to have their families in later life, only after having enjoyed a somewhat prolonged span of freedom from childbearing during the early years of marriage. In still other cases couples may distribute children rather widely over the years during which children may be born. In all three cases, the total number of children born may be the same--the difference being merely in the timing. If this timing is prescribed by social definition, or is a commonly accepted policy in a population at a given time, then the real cohorts that subscribe to a particular policy will have their children by a timing pattern that is different from that of cohorts that subscribe to other timing policies.

(5) Wars, epidemics, civil disorder, or other conditions may disrupt the life of family units by separating spouses or making it impossible for them to bear children. Such interruptions to conjugal life occur in particular years, and would have their greatest effect upon the real cohorts that are just nearing or are at the peak rate of marriage and family formation. These catastrophes may also affect the incidence of widowhood and divorce.

(6) Economic conditions may fluctuate a great deal. During periods of high level employment, easy credit, and high wages, families may be started more easily and at younger ages than under conditions of greater unemployment, low wages, and tightened credit. Hence, each cohort may bear the impress of the economic conditions it has experienced.

Thus, the fertility history of a real cohort is shaped by many factors. These factors may be cyclical or other temporary fluctuations or they may be fundamental changes in fertility attitudes and cultural prescriptions. These factors tend to impinge on each cohort in a unique pattern and sequence that is not exactly duplicated by any other cohort. The action of a cohort at any time is determined, not solely by conditions of the moment, but also by its cumulative experience up to the present and by the conditions which it anticipates to lie immediately ahead. Unlike mortality, which is largely involuntary and about which public attitudes, cultural practices, and private efforts all have a unilateral tendency, fertility is an event that is everywhere subject to some degree of human volition; hence, it is subject to influence by a much wider variety of forces. These forces do not all act uniformly in one direction; they may fluctuate in the direction of encouraging fertility at some moments and of discouraging it at others. Because of these facts, there is great need for a study of real cohort fertility or the study of fertility behavior over time.



18-7d. Definition of Real Cohorts. For most demographic analyses of fertility, real cohorts are defined as groups of persons born within specified intervals of time. These are called birth cohorts. They also could be defined in terms of marriage cohorts, or groups of persons who first married within specified intervals of time. Cohorts can also be specified by some other time-related event. The exposition which follows is in terms of birth cohorts, but the same principles can be applied to cohorts defined in other ways. Birth cohorts are identified by the year(s) of birth.

18-7e. Basic Cohort Fertility Rates. The same five basic fertility measures defined in cross-sectional terms in Section 18-3 may be defined in longitudinal terms as follows:

(1) Longitudinal Crude Birth Rate. If a real birth cohort (both sexes) were followed throughout its life, its crude birth rate would be defined as:

$$CBR(L) = \frac{B}{P} \cdot 1000 \quad (5)$$

where B is total births that occurred to the cohort during its lifetime, and

P is total number of person years of life lived by the cohort.

(2) Longitudinal General Fertility Rate is defined as:

$$GFR(L) = \frac{B}{{}_{30}C_{15}} \cdot 1000 \quad (6)$$

where  ${}_{30}C_{15}$  is the total number of person years of life lived by the female members of the cohort between their 15th and their 45th birthdays.

It is the average rate of childbearing (in terms of births per 1,000 women in the childbearing ages) for the female members during their reproductive span.

(3) Longitudinal Age-Specific Fertility Rate is defined as:

$$ASFR(L) = \frac{B_x}{C_x} \cdot 1000 \quad (7)$$

where  $B_x$  is the number of children born to female members of the cohort between their xth and (x+1)th birthdays, and

$C_x$  is the number of person years of life lived by the female members of the cohort between their xth and their (x+1)th birthdays.

These rates may be computed for single years of age or for 5-year or

other age groupings. They express the rate of childbearing by the cohort at each age, in terms of the births per 1,000 women of each age.

(4) Longitudinal Total Fertility Rate is defined as:

$$\text{TFR(L)} = \sum_{15}^{44} \text{ASFR(L)} \quad (8)$$

It is the average number of children 1,000 women of the cohort would bear in their lifetime if there had been no cohort mortality during the childbearing years.

(5) Age-Cumulative Fertility Rate is defined as:

$$\text{ACFR(L)} = \sum_{15}^x \text{ASFR(L)} \quad (9)$$

where x is some designated age.

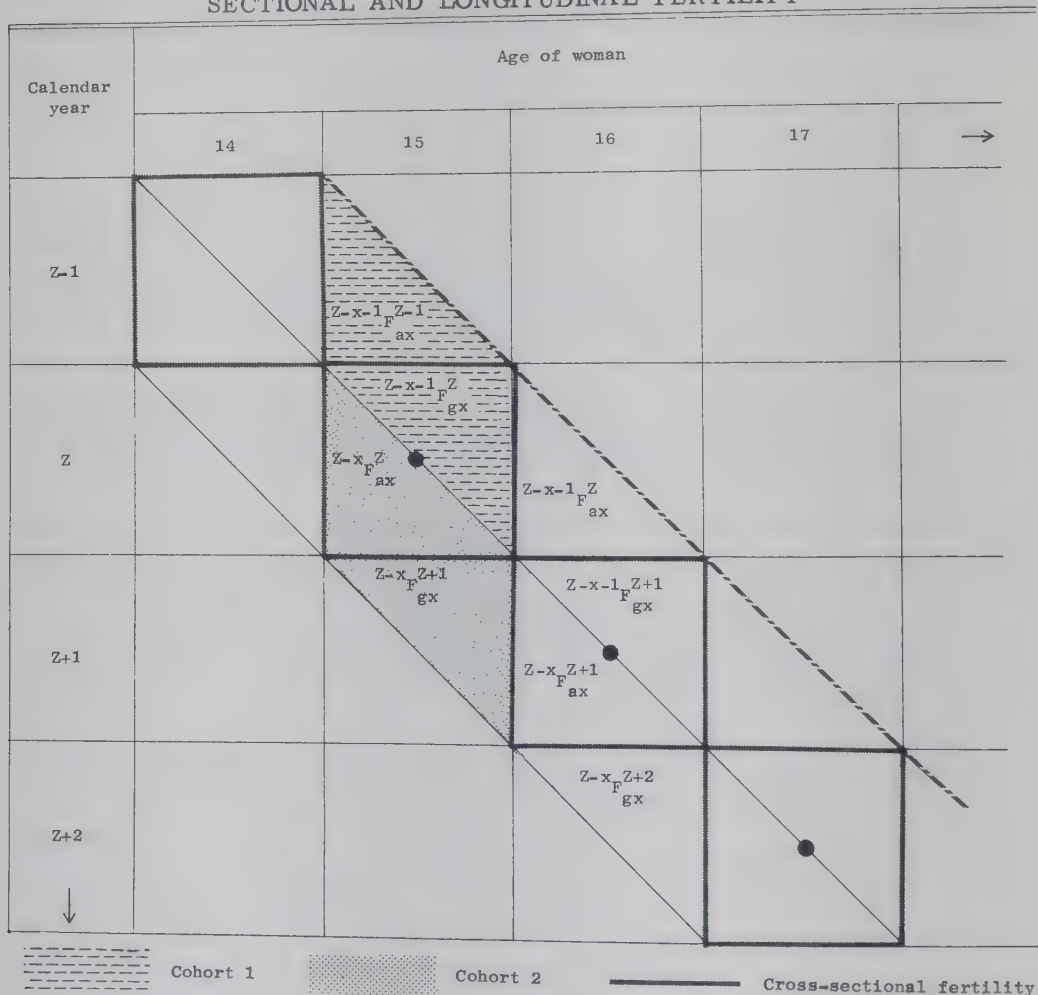
It is the average number of children 1000 women of the cohort would bear from the onset of childbearing up through the age specified by x, provided there was no mortality during the interval between age 15 and age x.

18-7f. Estimation of Births and Woman Years for Longitudinal Rates from Conventional Vital Statistics and Census Tabulations. The calculation of the basic fertility measures in cohort terms requires that births be allocated to women who were born in certain specified years and who were age x at the time of delivery, irrespective of the calendar year in which the birth occurred. Similarly, it requires the calculation of the number of person years the women spent in age x, irrespective of the calendar years in which the person-years at age x are spent. Unless the data for births and the census are both tabulated in terms of cohort identification (year of birth), it is impossible to tabulate any of the rates defined in Section 18-7e directly. Instead, it is necessary to estimate them from conventional calendar-year tabulations of births and population. The problem may be clarified by the diagram of Figure 18-2, which is a matrix with single years of age across the captions and single calendar years down the stub. This table represents the conventional form in which birth statistics are reported: each cell may represent the number of births to mothers of each particular age during a succession of calendar years. Alternatively, each cell may represent the average (mid-year) population of women of each particular age during a succession of calendar years.

The period measures of ASFR (Section 18-3) would be computed by dividing the number of births that are classified in a particular cell of the table by the number of women that are in the same cell of a

corresponding table referring to women. The mid-year population of the cell (indicated by a black dot), which is the denominator of the rates, lie along the principal diagonal of the table. The women who comprise this mid-year population are drawn from two real calendar year cohorts. Note, for example, the year of birth of women who bore a child at age 15 in year  $z$ . One part of them had celebrated their 15th birthday in year  $z-1$  and then had their child in year  $z$  before having their 16th birthday. Such girls must have had their 15th birthday during the interval between January 1 and December 31 of year  $z-1$ ; let us identify this as cohort 1. The second real cohort is comprised of

Figure 18-2. DIAGRAM REPRESENTING RELATIONSHIPS BETWEEN CROSS-SECTIONAL AND LONGITUDINAL FERTILITY



women who celebrated their 15th birthday sometime during year  $z$  and then had a child during the same year. Such girls must have had their 15th birthday during the interval between January 1, and December 31, of year  $z$ . Let us identify this as cohort 2. Let us next divide the women of each of these two cohorts into two groups. Those who had their birthday  $x$  in the year before bearing a child at age  $x$  will be identified by the symbol "a" before the age designation  $x$ .

Those who had their birthday  $x$  in the same year they bore a child at age  $x$  will be identified by the symbol "g" before the age designation.

Let us identify real cohorts by a superscript to the left and before the symbol identifying a population. For example, we could define the women aged  $x$  in 1970 the cohort born in 1940 as  $^{40}_{70}F_x$ . We can now identify the women who comprise the mid-year population of a given age in terms of the real cohorts (year of birth) to which they belong, as follows:

$^{Z-x-1}_ZF_{gx}$	Women aged $x$ in year $Z$ who had their birthday $x$ during the year preceding $Z$ (part of cohort 1)
$^{Z-x}_FZ_{ax}$	Women aged $x$ in year $Z$ who had their birthday $x$ during year $Z$ (part of cohort 2)

Figure 18-2 illustrates the real-cohort components of women aged 15 in calendar year  $z$ . If we identify births by age of mother, all of the births that occur in a given calendar year can be attributed to one of two cohorts as follows:

$^{Z-x-1}_BZ_{gx}$	Births to women aged $x$ in year $Z$ who had their birthday $x$ during the year preceding year $Z$ and who bore a child in year $Z$
$^{Z-x}_BZ_{ax}$	Births to women aged $x$ in year $Z$ who had their birthday $x$ during the year $Z$ and bore a child during year $Z$

By substituting these two sets of definitions into equation (3) we are able to define the ASFR for year  $Z$  in terms of the portions of real cohorts of which it is comprised.

$$\text{ASFR for women age } x = \frac{^{Z-x-1}_BZ_{gx} + ^{Z-x}_BZ_{ax}}{^{Z-x-1}_FZ_{gx} + ^{Z-x}_FZ_{ax}} \cdot 1000 \quad (10)$$

Study of Figure 18-2 will show that longitudinal age-specific fertility rates for a real calendar cohort (example for cohort 2) are computed as:

$$\text{ASFR(L)} = \frac{^{Z-x}_BZ_{ax} + ^{Z-x}_BZ_{gx}^{Z+1}}{^{Z-x}_FZ_{ax} + ^{Z-x}_FZ_{gx}^{Z+1}} \cdot 1000 \quad (11)$$

The dotted portion of Figure 18-2 identifies the groups that would be involved in computing a longitudinal age-specific fertility rate for women aged 15 in the cohort  $Z-x$  (cohort 2).

Equations (10) and (11) teach us that if we can divide the birth-days which women have into two components, one associated with  $ax$  and one associated with  $gx$ , we can compute ASFR either on a cross-

sectional or on a longitudinal basis, whichever we choose.

In order to divide conventional calendar year data into these two components, a separation factor of .5 is valid for all age groups.

We can now state a rule for estimating both the numerator and the denominator of real calendar year cohorts from conventional vital statistics tabulations of births and conventional mid-year populations of each age: Divide the data for each age group for each year by two. Then add the one-half for the year in which it first became possible for the cohort to have birthday x to the one-half for the same age group for the year following. Reading the above rule and noting the shaded portion of Figure 18-2 will make the procedure clear.

It is important for demographers to appreciate that this process of extracting longitudinal data from cross-sectional tabulations has the effect of smoothing year-to-year fluctuations by a moving average technique. The oft-noted smoothness of longitudinal measures is largely a function of the mode of deriving the data.

**IMPORTANT--**Please note that for real cohorts the "mid-year" population (population average age  $x + 0.5$  years) lies at the center of the diamond-shaped shaded area and that it has its  $x$ th birthday on an average at the end of year  $z$  and start of year  $z + 1$ . Thus a real cohort age  $x$  refers to women who are, on an average, one-half year younger than the period rate for the same age group.

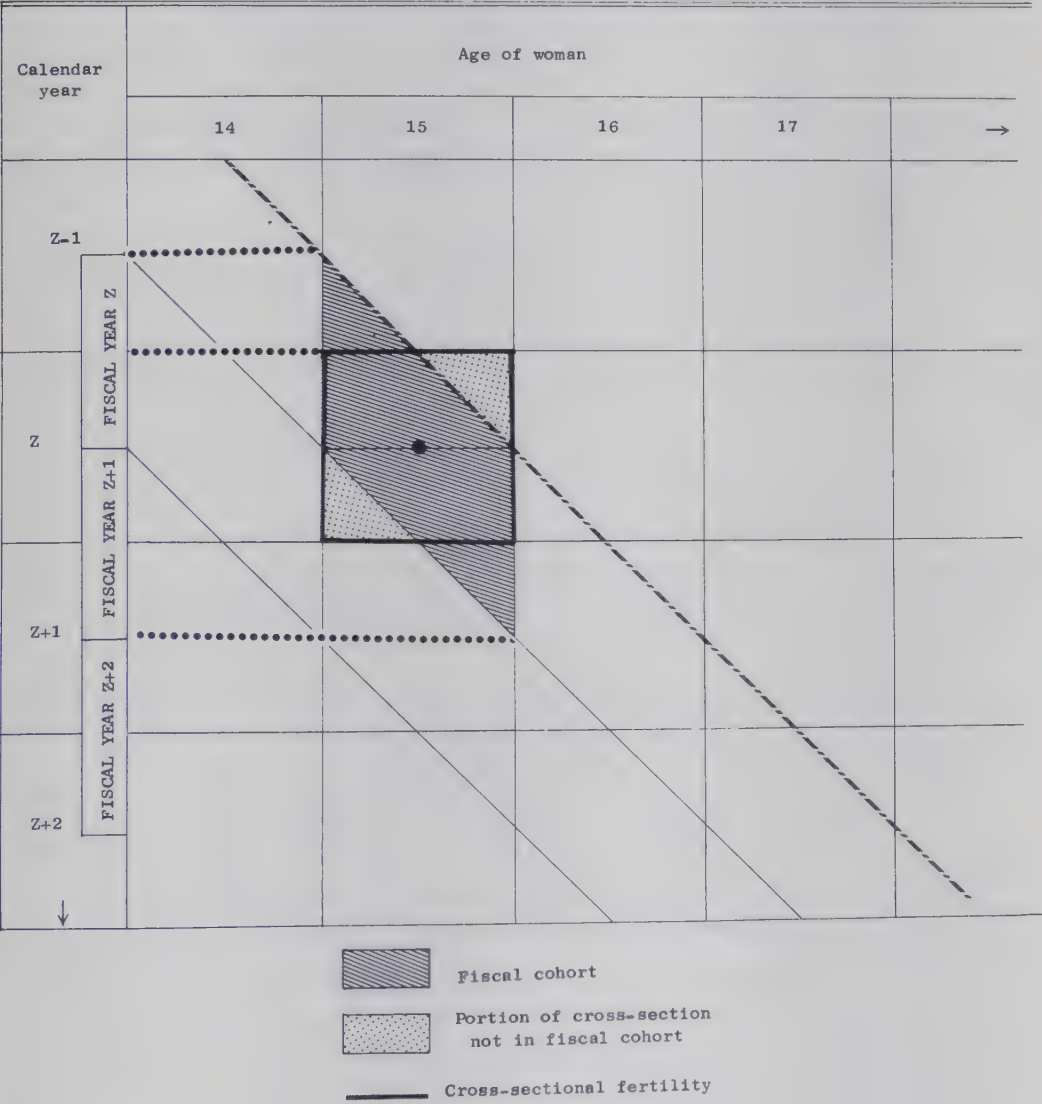
18-7g. Fiscal Real Cohorts. Figure 18-2 refers to real calendar year cohorts. An alternative procedure is to use fiscal cohorts, or groups of women born between July 1 of one year and July 1 of the next. (The term "fiscal year" is appropriate for the United States, where government fiscal accounts are kept in terms of years beginning July 1. We use the term here only to avoid cumbersome terminology.) Figure 18-3 illustrates the problem and its solution. It may be seen that the fiscal cohort contains women who were born in three calendar years ( $Z-1$ ,  $Z$ , and  $Z+1$  for age  $x$ ). The distribution of person years in year of age  $x$  for these women, however, is as follows:

<u>Year</u>	<u>Percent</u>
$Z-1$ .....	12.5
$Z$ .....	75.0
$Z+1$ .....	12.5
	<u>100.0</u>

The exact formula for deriving fiscal cohorts from conventional period tabulations is:

$$ASFR(L) = \frac{.125 B_x^{Z-1} + .750 B_x^Z + .125 B_x^{Z+1}}{.125 F_x^{Z-1} + .750 F_x^Z + .125 F_x^{Z+1}} \cdot 1000 \quad (12)$$

Figure 18-3. DIAGRAM REPRESENTING RELATIONSHIPS BETWEEN CROSS-SECTIONAL AND LONGITUDINAL FERTILITY BASED ON FISCAL YEARS



The fact that three-fourths of the women in a fiscal cohort have their birthday  $x$  in the same calendar year as they bear a child at age  $x$  leads naturally to an approximation of substituting the 25 percent of the births that occur to other cohorts at age  $x$  in year  $Z$  for the 12.5 percent from  $Z-1$  and 12.5 percent from  $Z+1$ . In other words, the conventional period ASFR is an approximation of the ASFR for a fiscal real cohort.

$$\text{Fiscal ASFR}(L) \div \text{ASFR} \quad (13)$$

Which of the two types of real cohorts should be used--the calendar year cohort or the fiscal cohort? It is recommended that unless considerations dictate otherwise, the fiscal cohort as defined in (13) be used for studies of longitudinal fertility. This procedure has the following advantages:

- (a) It is less of a moving average when calculated from period data than are calendar cohorts, and hence preserves more of the year-to-year uniqueness of cohort experience.
- (b) The average birthday is July 1, which is identical with ASFR; differences in age composition and time referent between real and hypothetical cohort measures are removed.
- (c) It is much easier to compute; since it is taken to be identical with ASFR, only one set of computations is needed.

When done in terms of single years of age and single calendar years, the error created by using the approximate equation (13) is so tiny as to be insignificant.

In its studies of longitudinal fertility from fertility histories, the U. S. Bureau of the Census has generally worked with calendar-year-of-birth or calendar-year-of-marriage cohorts.

18-7h. Age-Groupings of Real Cohort Data. The principle of adopting ASFR as an approximation of fiscal real cohort rates applies only to data for single years of age for single calendar years. Age-groupings for 5-year ages may be computed from the raw data for the single years, summing births and woman-years down the diagonal separately, and then dividing to get a rate. Longitudinal analysis usually is attempting to measure small differences, and any form of abridgement, except the precise one, introduces data for other cohorts into the analysis and tends to destroy the information that is uniquely derivable by the cohort approach.

18-7i. Example of Construction of a Real Cohort. In Principles of Demography,<sup>2</sup> age-specific fertility rates for single years of age by single calendar years from 1917 to 1965 are reported. From this table we may construct the fertility history of real cohorts. This is done by selecting a cohort of interest and following it down the diagonal of the table, in so doing, adding increments of one year of age and one calen-

dar year. Table 18-10 illustrates the fiscal approximation for this real-cohort ASFR for the cohort born in 1907.

Table 18-11 represents the complete schedule of ASFR for two very different cohorts: one which reached peak fertility during the midst of the economic depression of the 1930's and one that passed through its early years of high fertility during the "baby boom" that characterized United States fertility from 1946 to 1960. It is instructive to compare the schedules of age-specific fertility rates, the cumulative fertility rates, and the proportion distribution of fertility for these two cohorts. The "baby boom" cohort produced an average of one more child per woman than the "depression" cohort; yet after age 35 the ASFR for this cohort were lower than for the depression cohort. By the end of age 35, the baby boom cohort had borne 93.4 percent of its children, but by this age the depression cohort had borne only 87.9 percent. Apparently, the depression cohort was attempting to "make up" births that had been postponed during the economic depression.

Note that, for both cohorts, 50 percent or more of all childbearing is completed by age 25 and 90 percent by age 37. The more recent cohort is completing its childbearing at a significantly earlier age than the older cohort.

18-7j. Cohort Analysis of Fertility Change. If all fertility changes were due entirely to annual fluctuations in fertility levels, there would be no cohort effect. In this case, each fertility change that occurs would affect all cohorts equally. However, if some cohorts were changing their fertility more rapidly than other cohorts, we must acknowledge a cohort effect. Evidence of a cohort effect is provided by the indices of relative change in ASFR. In Table 18-7, for example, relative fertility decline is shown to be much greater for the younger than for the older ages. If all age groups were to show the same percent change as the GFR, there would be no cohort effect; significant deviations of relative change for particular ASFR from the relative change of GFR is evidence of a cohort effect in fertility change. It is possible to know not only the extent of overall cross-sectional change in fertility, but which particular cohorts are contributing to it. In some developing nations it has been discovered that when birth rates start to decline, it begins among the older cohorts of women who are weary of childbearing and spreads gradually to younger cohorts. In other situations, it may be the younger women in their late 20's and early 30's who, being less traditional and more willing to accept new ideas, initiate fertility decline. Modernization sometimes brings about a rise in fertility at the very youngest ages while promoting fertility decline at all other ages. In all cases, the measure of percent change in ASFR in comparison to percent change in GFR reveals the cohort differences.

Table 18-10. ILLUSTRATION OF PROCEDURE FOR CONSTRUCTING A REAL COHORT FROM CROSS-SECTIONAL TABULATIONS OF FERTILITY RATES FOR SINGLE YEARS OF AGE BY SINGLE CALENDAR YEARS: UNITED STATES COHORT BORN IN 1907

Calendar year	Age																			etc. to age 49		
	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32		33	34 →
1920.....	3.3																					
1921.....		7.1																				
1922.....			21.9																			
1923.....				52.8																		
1924.....					95.2																	
1925.....						127.1																
1926.....							136.0															
1927.....								136.7														
1928.....									142.1													
1929.....										145.0												
1930.....											137.1											
1931.....												126.4										
1932.....													115.4									
1933.....														111.0								
1934.....															110.0							
1935.....																96.5						
1936.....																	94.0					
1937.....																		77.9				
1938.....																			77.6			
1939.....																				69.5		
1940.....																					63.8	
1941.....																						
1942.....																						
etc. to 1957.																						

Source: Donald J. Bogue, Principles of Demography, Table 18-38, pp. 734-35.

Table 18-11. ILLUSTRATION OF PROCEDURE FOR CALCULATING TFR(L), CFR(L), AND MEASURES OF DISTRIBUTION OF CHILDBEARING FOR REAL COHORTS; UNITED STATES "DEPRESSION COHORT" BORN IN 1907 AND "BABY BOOM COHORT" BORN IN 1930

Age of cohort	"Depression cohort" born in 1907				"Baby boom" cohort born in 1938			
	Calendar year in age x	Age specific fertility	Age cumulative fertility ACFR (L)	Percent of total fertility by age x	Calendar year in age x	Age specific fertility	Age cumulative fertility ACFR (L)	Percent of total fertility by age x
14.....	1921	3.3	3	0.1	1944	3.5	3	0.1
15.....	1922	7.1	10	0.4	1945	8.9	12	0.4
16.....	1923	21.9	32	1.4	1946	23.6	36	1.1
17.....	1924	52.8	85	3.7	1947	66.6	103	3.2
18.....	1925	95.2	180	7.9	1948	116.4	219	6.7
19.....	1926	127.1	307	13.5	1949	158.4	377	11.6
20.....	1927	136.0	443	19.4	1950	179.3	557	17.2
21.....	1928	136.7	580	25.4	1951	108.7	765	23.6
22.....	1929	142.1	722	31.6	1952	220.6	986	30.4
23.....	1930	145.0	867	38.0	1953	231.5	1218	37.5
24.....	1931	137.1	1004	44.0	1954	231.8	1449	44.7
25.....	1932	126.4	1131	49.6	1955	220.0	1669	51.4
26.....	1933	115.4	1246	54.6	1956	209.7	1879	57.9
27.....	1934	111.0	1357	59.5	1957	196.1	2075	63.9
28.....	1935	110.0	1467	64.3	1958	177.6	2253	69.4
29.....	1936	96.5	1564	68.5	1959	161.9	2415	74.4
30.....	1937	94.0	1658	72.7	1960	145.4	2560	78.9
31.....	1938	77.9	1736	71.7	1961	125.9	2686	82.8
32.....	1939	77.6	1813	79.4	1962	108.9	2795	86.1
33.....	1940	69.5	1883	82.5	1963	91.6	2886	88.9
34.....	1941	63.8	1946	85.3	1964	79.4	2966	91.4
35.....	1942	60.9	2007	87.9	1965	64.2	3030	93.4
36.....	1943	58.2	2066	90.5	1966	49.3	3079	94.9
37.....	1944	51.3	2117	92.8	1967*	40.9	3120	96.1
38.....	1945	47.7	2164	94.8	1968	35.0	3155	97.2
39.....	1946	37.9	2202	96.5	1969	27.0	3182	98.0

Table 18-11. ILLUSTRATION OF PROCEDURE FOR CALCULATING TFR(L), CFR(L), AND MEASURES OF DISTRIBUTION OF CHILDBEARING FOR REAL COHORTS; UNITED STATES "DEPRESSION COHORT" BORN IN 1907 AND "BABY BOOM COHORT" BORN IN 1930 (CONTINUED)

Age of cohort	"Depression cohort" born in 1907					"Baby boom" cohort born in 1938				
	Calendar year in age x	Age specific fertility	Age cumulative fertility ACFR (L)	Percent of total fertility by age x	Calendar year in age x	Age specific fertility	Age cumulative fertility ACFR (L)	Percent of total fertility by age x		
40.....	1947	29.1	2232	97.8	1970	22.0	3204	98.7		
41.....	1948	18.2	2250	98.6	1971	15.2	3219	99.2		
42.....	1949	14.1	2264	99.2	1972	11.0	3230	99.5		
43.....	1950	8.5	2272	99.6	1973	7.0	3237	99.8		
44.....	1951	5.0	2277	99.8	1974	3.9	3241	99.9		
45.....	1952	2.8	2280	99.9	1975	2.2	3244	100.0		
46.....	1953	1.3	2281	99.9	1976	0.9	3244	100.0		
47.....	1954	0.5	2282	100.0	1977	0.2	3245	100.0		
48.....	1955	0.1	2282	100.0	1978	0.1	3245	100.0		
49.....	1956	0.1	2282	100.0	1979	0.0	3245	100.0		

Source: Donald J. Bogue, Principles of Demography, Table 18-38, pp 734-35, updated with data from annual reports Vital Statistics of the United States. For the "baby boom" cohort, the rates for ages 37 and above are the age-specific fertility rates for the year 1966. It is believed that these estimates very closely approximate the eventual fertility of this cohort and if biased tend to overestimate slightly the eventual fertility.

Although highly useful, the above procedure is only a first approximation to the study of cohort effects upon fertility change. More sophisticated approaches are only now in the process of development, and much research remains to be done. These approaches must begin with the recognition that fertility change is variation from some average or typical pattern, and that this variance must be partitioned into two components: one due to "annual effects" and one due to "cohort effects." If this is done, it is possible to understand better the year-to-year changes in fertility that occur. Bogue undertook to make such a division for the fertility changes in the United States in Section 18-14 of Principles of Demography. The procedure followed there was to:

- (a) Develop an average age-pattern of fertility by averaging cross-sectional ASFR for many years;
- (b) Derive the deviations of ASFR from this average for each age for each year--this is the "variation to be explained";
- (c) Estimate the cohort effect by isolating the unique age-patterns that emerge in particular years. This is done by setting up a model where TFR are equal for all calendar years but are distributed according to age by the relative size of ASFR to the actual TFR for the year. Subtracting average ASFR from such values reveals unique age effects that cannot be attributed to the average fertility level for the year;
- (d) Estimate the calendar effects as a residual, subtracting the deviations of (c) from the deviations of (b).

Results of these calculations for the United States revealed that for most years, the annual component is substantially larger than the cohort component. During times of major transition from high to low fertility, the cohort component tends to take on more importance; and during major transitions in marriage patterns it is very large. When averaged over half a century of experience, however, the annual component accounts for about twice as much variation as does the cohort component. If these findings can be generalized to other situations, it suggests that the major task in studying fertility change is to measure the annual change and then to attempt to use cohort analysis to modify or qualify the results of such cross-sectional analysis. Questions such as "Is the age-pattern of marriage changing?" and "Is the timing of childbearing within marriage changing?" need to be asked and answered, at least tentatively. In this way, meaning may be attached to the differences in relative fertility change observed for ASFR.

18-7k. Children Ever Born Data and Fertility Change. How useful are children ever born data taken from successive censuses in revealing fertility change? If they are used as cross-sectional indices of current fertility change, they can lead to inferences that are com-

pletely erroneous. When it is appreciated that CEB data refer to real cohorts and to cumulative (rather than current) fertility, one realizes that they must be used with caution and in the proper context. If the data are of equally good quality and were collected on a comparable basis, they can be most useful. The analysis may be conducted either on a straight cross-sectional basis, or it can be done by studying the changes in behavior of specific real cohorts in comparison to previous real cohorts when they passed through these ages. Table 18-12 illustrates the computations that may be made. The analysis may be made either in terms of overall ratios of children to women or in terms of proportions of women who have borne specific numbers of children of each order. If birth rates change in a population, especially at the younger ages, that fact will be signalled by age-specific rates of children ever born. For populations subject to fertility fluctuations, fertility change measurements computed from children ever born data must be interpreted with caution because they do not necessarily measure recent changes in fertility at the ages indicated. Because they are cumulative fertility measures, they are able only to state a net balance as of the date of the census. For example, we know that in 1960 the ASFR for women over 30 years of age were declining rapidly. Nevertheless, the CEB data for ages 30-39 show a great increase over 1950. This is due to the higher fertility of these women when they were younger, carried forward to older ages. Because of their cumulative nature, CEB data for low birth rate populations are insensitive to current fertility.

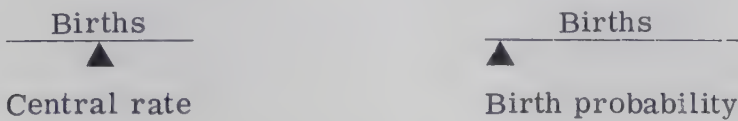
Table 18-12. USE OF CHILDREN-EVER-BORN DATA TO MEASURE FERTILITY CHANGE; FOR THE UNITED STATES: 1950-60

Age	Children ever born per 1000 women		Change	
	1960	1950	Absolute	Percent change
15-19 years.....	127	105	+ 22	+17.3
20-24 years.....	1030	738	+292	+28.3
25-29 years.....	2007	1436	+571	+28.5
30-34 years.....	2452	1871	+581	+23.7
35-39 years.....	2518	2061	+457	+18.1
40-44 years.....	2407	2170	+237	+ 9.8
45-49 years.....	2247	2292	- 45	- 2.0
50-54 years.....	2179	2497	-318	-14.6

As a general principle it must be acknowledged that there is no known way of estimating the level of current fertility from changes in the CEB data. The researcher is advised to make use of information on age distributions, on "own children" and from the regression equations to estimate measures of current fertility for two successive periods. (These procedures are described below. Whatever biases

each of these procedures may have, the same bias will tend to appear in the estimate for both dates, with the result that the measurement of change will tend to be surprisingly accurate.)

18-8. Central Rates Versus Birth Probabilities. All of the fertility rates presented above (both period and longitudinal) are central rates; they take as their base the mid-interval population. A birth probability, in contrast, takes as its base the population as of the beginning of the interval. It answers the question, "What is the probability that on her x-th birthday a woman will bear a child within the next 12 months?" The distinction between central rates and birth probabilities is clarified by the following diagrams:



This distinction was pointed out in the study of mortality, where the central death rate  $M_x$  was differentiated from  $q_x$ , the probability of dying in the year of age x. The appropriate date for the denominators for birth probabilities for the classes of rates presented thus far are:

ASFR (period)	January 1 of year Z
ASFR (Longitudinal-calendar)	July 1 of year Z
ASFR (Longitudinal-fiscal)	January 1 of year Z

These are the average birth dates of the cohorts involved.

The distinction between central fertility rates versus birth probabilities is therefore a matter of how one defines the women to which the denominators refer. Birth probabilities take as their base a population 6 months younger than the central rate for the same age. Births (the numerator) must then be tabulated or estimated to represent the fertility of the women while they are in the status specified in the denominator.

One unique difference between birth and death probabilities is that the act of giving birth does not affect the status of the mother as a member of the original cohort. Death removes the person from the cohort and obviously prevents her reappearance at later ages. The population of the cohort is changed and we must take this into account. Giving birth makes no such change necessary. The women who bear children in one year are eligible to be exposed to the likelihood of bearing children in the next.

On their x-th birthday a real cohort of women may be expected to be exposed during the next 12 months to two forces: fertility and mortality. The birth probability approach implicitly assumes there will be no mortality during the year. The central rates approach recognizes that mortality does occur throughout the year and that the mid-interval

population is an unbiased estimate of the women who are exposed to childbearing. However, it is possible to adjust the denominator of birth probabilities for mortality. If this is done it becomes a statement of the number of woman months that surviving women will spend between two successive birthdays. Because at younger ages fertility changes greatly within even six-month intervals, birth probabilities can have very different values from central rates.

18-9. Nuptial (Conjugal) and Illegitimacy Fertility Rates. A major criticism of the basic fertility measures is that they are not truly based on "exposed" populations. They include women who have never married or who are widowed or divorced; such women are not exposed to "legitimate" or socially normal childbearing. A refinement that is proposed, therefore, is to compute nuptial fertility rates, in which the numerators refer to legitimate births and the denominators refer to married women. Many nations of the world classify births as legitimate or illegitimate. In such cases, it is possible to calculate

$$\text{GFR}(N) = \frac{L_B}{M_F} \cdot 1000 \quad , \quad \text{and} \quad (14)$$

$$\text{ASFR}(N) = \frac{L_{B_x}}{M_{F_x}} \cdot 1000$$

where  $L_{B_x}$  are legitimate live births to women of age  $x$ , and  $M_{F_x}$  are married women of age  $x$ .

Table 18-13 illustrates the procedure for computing these rates.

It is important to understand the nature of the relationship between  $\text{ASFR}(N)$  and "regular"  $\text{ASFR}$ . If all births occur within marriage, we may write:

$$\text{ASFR} = \frac{B_x}{M_{F_x}} \cdot \frac{M_{F_x}}{T_{F_x}} \cdot 1000 \quad (15)$$

where  $M_{F_x}$  is the married female population of age  $x$ , and  $T_{F_x}$  is the total female population of age  $x$ .

But  $B_x/M_{F_x}$  is the definition of  $\text{ASFR}(N)$ ; and  $M_{F_x}/T_{F_x} = p(n)$ , or the proportion of women of age  $x$  who are married. Thus:

$$\text{ASFR} = \text{ASFR}(N) \cdot p(n) \quad , \quad \text{and} \quad (16)$$

$$\text{ASFR}(N) = \frac{\text{ASFR}}{p(n)} \quad . \quad (16a)$$

Thus, where all births are legitimate the nuptial fertility rates are simply "regular" fertility rates weighted by the proportion of women

Table 18-13. ILLUSTRATION OF PROCEDURE FOR CALCULATING NUPTIAL ASFR, AFR, GFR-STANDARDIZED, ACFR, AND TFR; UNITED STATES: 1960

Age	Married women (000)	Legitimate births (000)	ASFR(N) Age specific nuptial fertility rate	GFR(N) standardized (direct)		TFR(N) calculation		ACFR(N) Age cumulative fertility rate
				Standard population	Expected births	Weight	Weighted ASFR (N)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total.....	25,804.6	4,033.5	156.3	1,000.0	156.07	....	808.6	....
15-19 years.....	1,033.8	502.2	485.5	40.0	19.43	.246	119.5	598
20-24 years.....	3,834.0	1,358.9	354.4	146.8	52.66	.785	278.2	1,988
25-29 years.....	4,772.0	1,060.7	222.3	184.9	41.10	.949	211.0	3,043
30-34 years.....	5,423.2	668.8	123.3	210.2	25.92	.983	121.2	3,650
35-39 years.....	5,658.0	349.3	61.7	219.3	13.53	.993	61.3	3,956
40-44 years.....	5,083.6	93.6	17.4	197.0	3.43	1.000	17.4	4,043

ASFR(N) = column (2) divided by column (1) multiplied by 1,000  
GFR(N) = (4,033.5/25,804,600) 1,000 = 156.3  
GFR(N) directly standardized = (156.07/1,000) 1,000 = 156.07  
GFR(N) indirectly standardized = multiply a schedule of ASFR(N) taken as standard by the reported ever-married population of each age, to get expected nuptial births. Expected nuptial births divided by actual legitimate births yields adjustment factor that is multiplied by GFR(N) for the standard population (calculations not shown) to get indirectly standardized GFR(N)  
Weighted ASFR = column (3) multiplied by column (6). (Weights are derived from a nuptiality table; they specify the proportion of person-years spent in an ever-married status at each age by those surviving and married by age 40-45. Alternatively, they may be computed from census data by a procedure described in the text.  
TFR(N) = sum of weighted ASFR multiplied by age interval = (808.6)5 = 4,043. This implies that 1,000 ever-marrying female will bear 4,043 children if they marry according to age-pattern specified in column (6) and remain unmarried until end of the reproductive years meanwhile bearing children according to the schedule of ASFR(N) specified in column (3).  
CFR(N) (cumulative fertility rate) = weighted ASFR(N) multiplied by age interval and cumulated progressively by age groups = column (7) 5 cumulated.

who are married. We may use equation (16) to convert from one form to the other. This relationship holds for GFR as well as for ASFR.

When a significant proportion of births occur to women who are single, widowed, or divorced, the nuptial fertility concepts begin to take on moralistic, class and ethnic connotations which the researcher may not desire. Moreover, the fact that under these conditions only a portion of all fertility is included as nuptial fertility may be undesirable. Under these circumstances, it may be advisable to shift from the concept "de jure nuptial" (legally married) fertility to "de facto nuptial" (sexually active) fertility. In this case, all births are divided by a denominator which is an estimate of the female population of reproductive age who are cohabiting, irrespective of their legal marital status. From sample surveys or other estimates, a standard distribution of the proportion of women of each age who are cohabiting may be developed and used as a substitute for the distribution of legal marital status by age.

TFR(N) is defined as the number of children 1000 married women would bear if they were to marry according to a standard age pattern typical for that population, and thereafter were to remain married until the end of the childbearing ages. It is calculated by multiplying the ASFR(N) by a proportion which indicates the portion of those women who eventually marry before age 45 or 50 that actually have married by the time they attain each age between 15 and 45 or 50. Ideally, these proportions should be derived from a nuptiality table. Alternatively, they may be derived from the percent married at each age according to the most recent census. In this latter case, the percent ever-married at age 35-39, 40-44 or 45-49 (whichever is larger) is taken as a base and the percent ever-married for all other ages are divided by it. The result is a set of weights the highest value of which is 1.000. If age 35-39 or 40-44 is the base of the system, all ages following must be given a weight of 1.000 also. Table 18-13 illustrates the procedure and gives additional instructions.

Since nuptial TFR is a standardized rate, it is therefore subject to the limitations of all standardized rates. When comparing such rates where different schedules of nuptiality age distributions are employed, it is important to recognize that nuptiality is not held constant. To accomplish this, the same standard nuptial distribution must be used for all data being compared.

GFR(N) may easily be standardized for age composition, as illustrated in Table 18-13. In this example, the age distribution of the married population of the United States is used as the standard population. Indirect standardization of GFR(N) follows the logic for indirect standardization of regular rates. A note at the bottom of the table specifies the steps for indirect standardization of GFR(N).

Nuptial fertility rates can be calculated on a real cohort basis following the principles of Section 18-7.

Illegitimacy (Exnuptial) Rates. The women who are exposed to bearing illegitimate children are single, widowed, and divorced. It is possible to construct the same four measures (outlined above for nuptiality) simply by substituting illegitimate for legitimate births in the numerators and the "not-married women" for the "married women" in the denominators. Christopher Tietze has suggested that these rates should be called "exnuptial rates" rather than "illegitimacy rates." We favor this change, but we will also retain the old terminology in the exposition which follows.

1. General illegitimate (exnuptial) fertility rate:

$$\text{GFR}(E) = \frac{{}^i_B}{N_F} \cdot 1000 \quad (17)$$

where  ${}^i_B$  is the number of illegitimate (exnuptial) births, and  $N_F$  is the number of single, widowed or divorced females of childbearing age (not-married female population).

2. Age-specific illegitimate (exnuptial) fertility rates: age-specific illegitimate (exnuptial) rates are a straightforward extension of definitions already established.

$$\text{ASFR}(E) = \frac{{}^i_{B_x}}{N_{F_x}} \cdot 1000 \quad (18)$$

GFR(E) will be greatly affected by age distribution of the not-married population and, therefore, should be standardized by age. The procedure is illustrated in columns (4) and (5) of Table 18-14. A representative age distribution of 1000 not-married women is established and is used as a standard to weight the ASFR(E). Because in this example the standard population is derived from the example itself, GFR(E) standardized and unstandardized are equal.

For TFR(E) it is necessary only to sum ASFR(E) and multiply by 5. This indicates the number of illegitimate children 1000 women would bear if they were to remain unmarried throughout their reproductive lives and, meanwhile, be exposed to childbearing at each age according to the rates specified by ASFR(E).

The illegitimacy (exnuptial) rates can also be calculated on a real cohort basis by following the principles of Section 18-7 in making the calculations.

18-10. Calculation of CBR and GFR from Age Distributions. As yet many nations do not have adequate tabulations of statistics for births from which to measure fertility rates. In many, if not most of these same nations, however, a reasonably accurate census is now

Table 18-14. ILLUSTRATION OF PROCEDURE FOR CALCULATING ILLEGITIMACY (EXNUPTIAL) ASFR, GFR, GFR-STANDARDIZED, ACFR, AND TFR; UNITED STATES: 1960

Age	Unmarried women (000) (1)	Illegitimate births (000) (2)	ASFR(E) age specific illegitimate birth rate (3)	GFR (E) standardized		ACFR(E) cumulative fertility rate (6)
				Standard population (4)	Expected births (5)	
Total.....	10,288.8	219.7	21.4	1000	21.40	....
Age 14 years — 1/	1,330.9	4.6	3.5	....	0.46	18
15-19 years.....	5,554.8	87.1	15.7	540	8.91	96
20-24 years.....	1,686.0	68.0	40.3	164	6.61	298
25-29 years.....	765.1	32.1	42.0	74	3.11	508
30-34 years.....	688.2	18.9	21.7	67	1.45	616
35-39 years.....	760.5	10.6	13.9	74	1.03	686
40-44 years.....	834.2	3.0	3.6	81	0.29	704

1.  $GFR(E) = (219.7/10,288,800) 1,000 = 21.4$ .
2.  $ASFR(E) = \text{column (2) divided by column (1) multiplied by } 1,000$ .
3. Expected births = column (3) multiplied by column (4).
4.  $GFR(E)$  directly standardized = sum of expected births divided by 1,000 (total of standard population).
5.  $TFR(E) = \text{sum of column (3) multiplied by age interval} = (14.07)5 = 704$ .
6.  $CFR(E) = \text{cumulative/fertility} = \text{cumulation of } ASFR(E) \text{ multiplied by } 5$ .
7.  $GFR(E)$  indirectly standardized = multiply a schedule of  $ASFR(E)$  taken as standard by the reported never-married population of each age to get expected exnuptial births. Expected exnuptial births divided by actual exnuptial births yields an adjustment factor that is multiplied by  $GFR(E)$  for the standard population to get  $GFR(E)$  indirectly standardized (calculations not shown).

available. Almost invariably, these censuses provide detailed tabulations of the population by age and sex. By techniques to be described in this section it is possible to compute the crude birth rate and the general fertility rate from these age-sex distributions. An approximation to age-specific, total, and cumulative fertility rates can also be made. Thus, it is possible to compute or estimate from census age data all of the five basic fertility measures developed in Section 18-3. Clearly, this should have great utility in estimating fertility in nations where vital registration is defective. It means that at least once each decade a measurement can be made. The applicability of these techniques certainly is not confined to the nations that lack accurate and complete counts of vital events, however. Vital statistics necessary for conducting studies of trends over earlier decades may be completely lacking or too inaccurate to be of use. Even where highly accurate registration systems are operating, the birth data may fail to provide information for the particular units of area that are needed to make a particular study. For example, births may not be tabulated for census tracts or for individual cities and small towns within the state or province or county. With the techniques developed below, fertility measures may be generated for any area for which age-sex data have been tabulated. Moreover, these measures may be standardized to control area-to-area differences in age.

If a census is taken in a given year  $z$ , it is almost self-evident that the children enumerated as being under one year of age (in the first year of life) were born during the 12 months preceding the date of the census, or between the census date in year  $z$  and the same date in year  $z-1$ . If the population is a closed one (no in-migration or out-migration), the infants enumerated must be the survivors of the infants that actually were born during that year. By multiplying the census count of children under one year of age by a reverse survival ratio, the number of births which the children represent can easily be determined.

The reverse survival ratio is obtained in the following manner. In a life table population, the number of children living in their first year of life in a continuously born cohort is specified by  $L_0$ . These represent the survivors of births  $l_0$ , which is usually taken as 100,000. Thus, the proportion who have survived from birth and are in the year of age 0-1 is specified by the ratio  $L_0/l_0$ . Conversely, the ratio of the original cohort of births to the survivors is the reciprocal of this, or  $l_0/L_0$ . This ratio, multiplied by the census count of persons under one year of age (adjusted for underenumeration), gives the estimated births during the calendar year preceding the census.

$$\text{Births in year } z-1 = \frac{P_0^Z}{C_0} \cdot \frac{l_0}{L_0} \quad \text{or} \quad \frac{P_0^Z}{{}_0S_B \cdot C_0} \quad (19)$$

where  $P_0^Z$  is the population aged 0-1 years at census count in year  $z$ ,  
 $L_0$  and  $l_0$  are life table values,  
 ${}_0S_B$  equals  $L_0/l_0$ , the estimated proportion of total births in year  $z-1$  that survive to be in age 0-1 at the time of the census, and  
 $C_0$  is the proportion of children aged 0-1 years enumerated by the census. This is a correction factor which converts the census count of persons in an age group (0-1 years) to an estimate of the "true count" of this age.

The expression  ${}_0S_B$  is a conventional survival ratio. Using it as a divisor (multiplying by its reciprocal) converts it into a reverse survival ratio, or a ratio which converts survivors into estimates of number of living persons at some earlier date.

This method presumes that a life table is available which specifies a mortality experience that is similar to that of the population at the date for which there is a census. Even where the life table must be borrowed from another nation with similar mortality experience or derived from a model life table, it has been found that the resulting estimates of births can be sufficiently accurate to lead to reasonably correct estimates of the birth rate. This does not imply that just any life table may be used. It is essential that whatever life table is used, it be the best possible estimate of the force of mortality (especially at the younger ages) actually at work in the population during the years for which a fertility estimate is being prepared.

The degree of precision that will be demanded in any specific case will be determined by how much knowledge is already available. If a nation knows nothing of its fertility level, the demographer will be performing a genuinely important service if his estimates are correct within 5 or 10 percent. If census data are carefully adjusted for errors, results can be obtained that are quite precise.

By a line of reasoning similar to that laid down for equation (19) it may be seen that children enumerated as being in the age range of 1-2 years at the census date in year  $z$  (those aged one year at the last birthday before the census) were born during the period exactly 12-24 months preceding the date of the census. Applying the above equation, modified by the change in year of age, yields;

$$\text{Births in year } z-2 = \frac{P_1^Z}{C_1} \cdot \frac{l_0}{L_1} \quad \text{or} \quad \frac{P_1^Z}{{}_1S_B \cdot C_1}$$

Using the following general equation, an entire age distribution can be converted to estimates of births in preceding years:

$$\text{Births in year } z-x-1 = \frac{P_x^z}{C_x} \cdot \frac{1_0}{L_x} \quad \text{or} \quad \frac{P_x^z}{{}_xS_B \cdot C_x} \quad (20)$$

where  ${}_xS_B$  measures survival from birth to be in year of age  $x$ ,  
and  
 $C_x$  is the net undercount of persons age  $x$ .

If there has been an overcount at any age,  $C_x$  may be greater than 1.0.

These estimates must be made, of course, from tabulations by single years of age. These single year of age data may be interpolated from 5-year age distributions or smoothed to reduce errors due to apparent misstatement of age. (Such corrections may be reflected in the values of  $C_x$  employed.) If a life table is available for both the beginning and end of the decade, interpolation between life table values should be made to obtain values of  ${}_xS_B$ .

The procedure for converting census age distributions into estimates of live births is illustrated in Table 18-15 for the United States in the decade 1950-60. The procedure follows that described above. Once the births have been estimated by 12-month intervals preceding the census (column 7) it is useful to convert them to a calendar year basis. This is done by shifting from a date-of-census basis to a calendar year basis. Since the census is taken on April 1 in the United States, this involves taking three-fourths of the births for  $z-x$  and adding to it one-fourth of the births for age  $z-x-1$ . The last two columns of Table 18-15 compare the estimated live births with the official count, adjusted for underregistration.

The results of Table 18-15 demonstrate the validity of the procedure. The number of births for the 1950-60 decade, estimated by reverse survival of the census age distribution, differs from the official estimate of registered births (adjusted for underregistration) by only one percent. The overall estimate of 98 percent completeness of enumeration at all ages is a very rough correction; nevertheless it pulls the data into surprisingly close correspondence with what appears to be the demographic fact.

By definition, the estimates of births obtained in this manner can have three sources of error. First is the underenumeration of children. Second is the tendency to misstate the ages of young children at the time of the census. Third is the validity of the reverse survival ratio employed. Since infant mortality is much larger than mortality at ages 2-10; and since child underenumeration is usually greatest for infants under 1 year of age, estimates of fertility based upon census counts of infants under 1 year of age will probably be most seriously in error. More satisfactory estimates will be obtained with data for ages 1-10 inclusive at the time of the census. As pointed out below, it will usually be a good procedure to group the data for several years to obtain an average fertility measure for 5 to 10 years.

Table 18-15. WORKSHEET FOR CONVERTING CENSUS AGE DISTRIBUTION FOR CHILDREN INTO ESTIMATES OF LIVE BIRTHS FOR CALENDAR YEARS PRECEDING THE CENSUS; UNITED STATES: 1960 CENSUS

Age at census date	Children (000)			Survival ratios $S_x B$		Estimated births (000)				Actual births	
	$P_x$ reported at census	$C_x$ estimated proportion enumerated	Estimated true number	1960	1950	Inter-population $Z-x$	12-month intervals prior to census	Calendar years	Births in calendar years	Number (000)	Ratio estimated/actual
Under 1 year.....	4,112	.980	4,196	.97815	.97429	.97815	4,290	1959	4,294	4,295	.9998
1 year.....	4,106	.980	4,190	.97324	.96913	.97283	4,307	1958	4,307	4,255	1.0122
2 years.....	4,099	.980	4,183	.97192	.96734	.97100	4,308	1957	4,288	4,308	.9954
3 years.....	4,016	.980	4,098	.97102	.96616	.96956	4,227	1956	4,221	4,218	1.0001
4 years.....	3,988	.980	4,069	.97031	.96523	.96828	4,202	1955	4,195	4,104	1.0222
5 years.....	3,954	.980	4,035	.96969	.96445	.96707	4,172	1954	4,138	4,078	1.0147
6 years.....	3,820	.980	3,898	.96916	.96375	.96591	4,036	1953	4,028	3,965	1.0159
7 years.....	3,787	.980	3,864	.96868	.96313	.96480	4,004	1952	3,969	3,913	1.0143
8 years.....	3,649	.980	3,723	.96824	.96256	.96370	3,863	1951	3,820	3,823	.9992
9 years.....	3,482	.980	3,553	.96784	.96203	.96261	3,691	1950	3,692	3,632	1.0165
10 years.....	3,481	.980	3,552	.96747	.96152	.96152	3,694	....	.....	.....	.....
0-4 years.....	20,321	....	.....	.....	.....	.....	.....	1955-60	21,305	21,180	1.0059
5-9 years.....	18,692	....	.....	.....	.....	.....	.....	1950-54	19,647	19,411	1.0122

Source: Children by single years of age reported in 1960 Census of Population, Vol. I, Characteristics of the Population, Part 1, U.S. Summary, Table 46. Estimated completeness of enumeration taken from the same source: "Considering the evidence now available, a reasonable estimate of the rate of net undercount in 1960 seems to be about 2.0 percent....", page xvii. Survival ratios taken from official life tables of the U.S., 1949-51 and 1959-61, published by U.S. National Center for Health Statistics. Estimate of births for calendar years obtained by taking 75 percent of births in 12-month interval prior to census and adding 25 percent of births preceding 12-month interval. The official report of births are as reported in U.S. National Center for Health Statistics, Vital Statistics of the United States, Table 1-2. These figures have been adjusted for estimated underregistration.

Calculation of CBR. With the total number of births for each of the years preceding a census thus established, it is a routine matter to calculate the crude birth rate. We obtain an estimate of the total population at the midpoint of each of the years preceding the census date (using the date of the census as the beginning of a year) and divide the estimated births by the estimated total population. Since these years refer to time before the census, one of the easiest and most reliable ways to obtain an estimate of the total population is to interpolate between two censuses to get the appropriate total midyear population that corresponds to the years for which an estimate of births has been made.

Table 18-16 illustrates this interpolation procedure. Since the live births have been estimated for calendar years, the interpolation is for successive July first dates for each intercensal year. Because the estimate of births for a single year may contain errors that reflect age misstatement, a good procedure is to sum estimated number of births for five years, sum the estimated population for the same years and divide the former by the latter to obtain a five-year average CBR. If estimates of the annual number of births in each year between two censuses are thus divided into two parts, it is possible to detect a change in the birth rate during the decade.

From Table 18-16 it may be seen that the estimate of CBR derived from the census age distributions differs from the official figure by only about two percent. The estimated total population was inflated by 2.0 percent for net undercount, as was the age distribution for children. The official United States rates do not incorporate this correction. Had this inflation not been made, the estimated and the official rates would have corresponded almost exactly when taken in terms of either single years or five-year average periods. This demonstration that it is possible to estimate the crude birth rate as being 24.5 for the decade while the "official" rate is 24.9, illustrates the potential utility of this procedure.

Calculation of GFR. In order to calculate the general fertility rate, we need an estimate of the number of women aged 15-44 for each year for which an estimate of births was made. This can be obtained in at least three ways:

- (a) Interpolate between two censuses in the manner suggested above for estimating total population, using the number of women 15-44 at each census.
- (b) Interpolate on a cohort basis by averaging the population 10-39 at the first census and the population 20-49 at the second census ten years later to get the population age 15-44 at the mid-decade.
- (c) Use the same reverse survival ratio procedure employed to

RS-130  
12349 N71

estimate births to estimate the number of women 15-44 in each year.

$${}_{30}F_{15}^{z-x-1} = \frac{F_{15+x+1}}{{}_{30}C_{15}} \cdot \frac{T_{15+x} - T_{45+x}}{T_{15+x+1} - T_{45+x+1}} \quad (21)$$

where  ${}_{30}C_{15}$  is an estimate of the percentage of women aged 15-44 who were enumerated in the census,  
 $T$  refers to the cumulative stationary population of an appropriate life table for females,  
 $F$  refers to census counts of females, and  
 $x$  refers to age of children being reverse-survived.

The right hand term is the reverse survival ratio for women; it converts surviving females into estimates of the number of women 15-44 years of age who were alive during the  $z-x-1$ th year preceding the census.

Putting numerator and denominator together into a single expression, the formula for computing GFR from a census age distribution is:

$$GFR = \frac{\frac{P_x^{z-1} l_0}{C_x L_x}}{\frac{F_{15+x+1}}{{}_{30}C_{15}} \cdot \frac{T_{15+x} - T_{45+x}}{T_{15+x+1} - T_{45+x+1}}} \quad (22)$$

where  $F$  refers to females,  
 $P$  refers to total population (children),  
 $x$  refers to age (of children),  
 $z$  refers to the original year in which the census was taken,  
 $l_0$  is the radix of a relevant life table, and  
 $T_x$  (and other values of  $T$ ) refer to the cumulative stationary population of a relevant life table.

As we recommended for the CBR, the estimates for 3 or 5 years should be grouped to obtain an average 3-year or 5-year estimate of GFR. If all years for an intercensal period are divided into two parts and an average is computed for each part, changes in fertility may be measured if the data are regarded as sufficiently precise.

Table 18-16 is a worksheet illustrating the above computations for each year of the intercensal period 1950 to 1960 for the United States. The estimates computed in this way are compared with the rates as actually published, with an indication of the estimate of relative error. As we found for CBR, the estimates derived from the age distributions are somewhat lower than the official estimates adjusted for underregistration. This seems to result from the fact that the same adjustment of 2.0 percent for net undercount was applied both to children and to adults, whereas the official data are not adjusted for undercount of adults. Nevertheless, even as the estimates stand they are impressively precise. The official data indicate an increase of 7.6 points (6.7 percent) in GFR between 1950-54 and 1955-59. The

Table 18-16. WORKSHEET FOR CALCULATING GFR AND CBR FOR A DECADE PRECEDING A CENSUS  
UTILIZING ESTIMATES OF BIRTHS DERIVED FROM CENSUS AGE DISTRIBUTIONS

Calendar year	Estimated population as of July 1 (000)		Estimated births from age distribution	General fertility rate			Crude birth rate		
	Women aged 15-44	Total population		Estimated	Official	Ratio est/off	Estimated	Official	Ratio est/off
1959.....	36,603	180,750	4,294	117.3	120.2	.976	23.8	24.3	.979
1958.....	36,308	177,703	4,307	118.6	120.2	.987	24.2	24.5	.988
1957.....	35,923	174,681	4,288	119.4	122.9	.972	24.5	25.3	.968
1956.....	35,672	171,518	4,221	118.3	121.2	.976	24.6	25.2	.976
1955.....	35,484	168,438	4,195	118.2	118.5	.997	24.9	25.0	.996
1954.....	35,378	165,188	4,138	117.0	118.1	.991	25.1	25.3	.992
1953.....	35,285	162,200	4,028	114.2	115.2	.991	24.8	25.1	.988
1952.....	35,207	159,585	3,969	112.7	113.9	.989	24.9	25.1	.992
1951.....	35,116	157,124	3,820	108.8	111.5	.976	24.3	24.9	.976
1950.....	35,042	154,967	3,692	105.4	106.2	.992	23.8	24.1	.988
1955-59.....	179,990	873,090	21,305	118.4	120.6	.981	24.4	24.9	.980
1950-54.....	176,028	799,064	19,647	111.6	113.0	.988	24.6	24.9	.988

Source: Estimated population are official census figures interpolated to July 1, inflated by 2.0 percent for estimated underenumeration. Interpolated figures are reported in U.S. Bureau of Census, Current Population Reports: Population Estimates, Series P-25, No. 310, 1965. Official fertility and birth rates are as reported in U.S. National Center for Health Statistics, Vital Statistics of the United States, Table 1-2. The rates have been adjusted for estimated underregistration.

estimates derived from the age distribution indicate an increase of 6.8 points (6.1 percent). Where errors of age reporting are great ('heaping' on particular digits), this level of precision will not be attained. Nevertheless, with careful adjustment and smoothing of the data, significant changes in fertility should be detected even where the data are quite rough. This technique will be found to be especially informative where the procedure is carried out for two or more consecutive censuses.

The reader should note that if the estimate of underenumeration of children ( $C_x$ ) is deemed to be equal to the estimate of underenumeration of adult females ( ${}_{30}C_{15}$ ), the two correction factors cancel each other and both may be dropped. A serious bias could result if we made a correction factor for children without making one for adult females too.

If estimates for five-year periods are to be made directly from 5-year age groupings of the age distribution, the following equations will give a close approximation.

$$\text{GFR for 5 years preceding census} = \left( \frac{{}_5P_0}{{}_5C_0} \cdot \frac{{}_5I_0}{T_0 - T_5} \right) \div \left( \frac{{}_{30}F_{15}}{{}_{30}C_{15}} \cdot \frac{{}_5T_{20} - {}_5T_{50}}{{}_5T_{15} - {}_5T_{45}} \right) \quad (2)$$

$$\text{GFR for 5-10 years preceding census} = \left( \frac{{}_5P_5}{{}_5C_5} \cdot \frac{{}_5I_0}{T_5 - T_{10}} \right) \div \left( \frac{{}_{30}F_{20}}{{}_{30}C_{20}} \cdot \frac{{}_5T_{25} - {}_5T_{55}}{{}_5T_{20} - {}_5T_{50}} \right) \quad (2)$$

Calculation of ASFR, TFR, and CFR. An approximate schedule of ASFR for a population can be established from its census age distribution. The procedure follows.

1. Select a schedule of ASFR that is regarded as a "pattern" of the unknown schedule of ASFR for the population under study. By "pattern" we mean that the ratios between the successive ASFR in the selected schedule are thought to be identical or very similar to those in the ASFR for the population under study.
2. Multiply each rate in the schedule of "pattern" ASFR by the number of women of corresponding age in the population under study, to obtain an estimate of the number of children that would be born to mothers of each age in the population under study, if the "pattern" ASFR were to apply to them. If desired, this number of women in each age group may be a five-year average, derived as described above.
3. Sum the number of births for all age groups obtained in Step 2. This is the total number of births that would be expected in a population, given its actual age distribution and the assumption that the pattern ASFR applied to it.
4. Calculate an adjustment factor by taking the ratio of the

"actual" number (estimated by reverse survival of age distributions) to the estimated number obtained in Step 3.

5. Multiply each ASFR in the "pattern" schedule of rates by this correction factor. The result is a set of ASFR that is consistent with the total number of births estimated.

The validity of this procedure rests on the assumption that the pattern distribution of ASFR is representative of the actual distribution and that differences between the two are proportionately distributed among the various age groups. The major consideration is to select a pattern that reflects the age-at-marriage situation characteristic of the population and one that has a fertility level reasonably close to the one to be estimated. Table 18-17 illustrates the above procedure for calculating ASFR and TFR from census age distributions.

A variety of procedures may be used to select the pattern ASFR. One procedure is to use the schedule of ASFR for the nation or some other population thought to be similar. A better procedure for establishing ASFR is to use the regression equations to be developed in Section 18-17.

With the procedures suggested in this section, it is possible to generate estimates of the five basic fertility measures for any nation in the world or any part thereof for which age-sex data have been tabulated. If the data are very poor, an estimate based on that segment of the age distribution thought to be most reliable may be used. Even though the estimates may be crude, they will be far more informative than unreliable vital statistics. As each new census is taken, revised estimates should be made.

Table 18-17. ESTIMATION OF AGE-SPECIFIC FERTILITY RATES FOR ENGLAND AND WALES, USING DATA FOR FRANCE AS A PATTERN: 1960-61

Age	ASFR for France, 1960	Women of childbearing age, England and Wales (000)	Expected births (000)	Estimated ASFR for England and Wales	Official ASFR for England and Wales	Ratio estimated to official
<b>General Fertility Rate</b>	<b>83.9</b>	<b>10705</b>	<b>824.9</b>	<b>...</b>	<b>798</b>	<b>...</b>
15-19 years.....	21.7	1509	32.7	20.7	34.2	.60
20-24 years.....	168.6	1451	244.6	160.4	165.5	.97
25-29 years.....	181.6	1414	256.8	172.8	171.8	1.01
30-34 years.....	112.1	1497	167.8	106.7	100.6	1.06
35-39 years.....	53.2	1680	89.4	50.6	46.2	1.10
40-44 years.....	20.3	1499	30.4	19.3	13.8	1.40
45-49 years.....	1.8	1665	3.0	1.7	0.8	2.12

Actual number of births = 785,005  
TFR (estimated) = 2661

Ratio estimated to actual births = .95164  
TFR (official) = 2664

Ratio of estimated to official TFR = .999

It should be pointed out that the procedures outlined above may be applied equally well to age distributions derived from sample surveys of the population.

18-11. Child-Woman Ratios and their Relationship to the Basic Fertility Rates. For many years, demographers have employed the child-woman ratio (number of children per 1000 women of childbearing age) as an index of fertility. This index can be computed quickly from age-sex distributions of census data, by the following formula:

$$CWR = \frac{P_x}{F} \cdot 1000 \quad (25)$$

where  $P_x$  is an age grouping of children, and  
 $F$  refers to women of childbearing age.

Various ages of children and women have been used to compute this ratio. The most common are:

<u>Children 0-4</u>	<u>Children 5-9</u>	<u>Children 0-9</u>
<u>Women 15-49</u>	<u>Women 20-54</u>	<u>Women 15-54</u>

The similarity between this measure and the GFR is immediately apparent. If we compare equation (25) with equation (23) of Section 18-10, we see that the CWR and GFR differ only in the following ways:

- (a) the CWR makes no allowance for mortality either to children or women;
- (b) the CWR makes no correction for underenumeration and age misstatement of both children and women;
- (c) the CWR uses only an approximate rather than a precise specification of the population that bore the children contained in the numerator.

When the CWR is employed as a fertility index, it is necessary to assume that all three of the above differences are equal in the populations being compared, and hence are constants in the analysis. This assumption may be justified when comparing groups within a nation for a particular census, but is difficult to justify for comparing populations of different nations or populations at different censuses within a nation where the level of infant mortality has changed. An additional difficulty with the use of CWR is that it has no inherent theoretical meaning; it is an index but not a rate of fertility. It does not "fit" meaningfully into the basic system described in Section 18-3. It cannot be compared with GFR, which it resembles, for populations where GFR has been calculated.

There is a high correlation between CWR and the basic measures of fertility, as is shown by the following statistics (based upon the same 50 nations with reliable data, reported in Section 18-5):

<u>Basic fertility measure</u>	<u>Correlation with child-woman ratio</u>
Crude birth rate .....	.930
General fertility rate.....	.961
Total fertility rate .....	.964

It is evident that the correlations are high, but it must be pointed out that they are not nearly so high as the interrelationships among the basic fertility measures themselves (especially GFR and TFR), as reported in Section 18-5. Therefore, the CWR method of measuring fertility is substantially less precise than the use of GFR, as estimated in Section 18-10.

The data needed to calculate the CWR are identical to those needed to calculate the GFR from census data, as described in Section 18-10. Equations (25) and (23) are therefore alternative modes of procedure. The virtue of the CWR is greater ease of calculation. The use of CWR is not recommended. Some demographers favor its retention because they regard it as a measure of "effective fertility;" this is fertility after infant and early childhood mortality has exerted its effect.

However, as Section 18-17 will demonstrate, the CWR is very useful as a key item of the raw material needed to generate with fairly high precision all of the basic fertility measures from census data using multiple regression equations.

18-12. Estimation of Basic Fertility Measures from Children-Ever-Born Data: Cumulative Longitudinal Fertility. In an effort to increase knowledge of fertility, national censuses have included special fertility questions in their enumerations. The question most frequently asked, and the one that has proved most useful is the children-ever-born question. Each ever-married woman is asked, "How many live-born children have you had? (Be sure to include children that may have lived only a short while and children that may have been in a previous union.)" If conscientiously asked of even a small but representative fraction of the population and properly recorded, coded, and tabulated, the information obtained is highly useful. Techniques for manipulation of these data to extract basic fertility measures are similar to (but not identical with) those already discussed. There has been a great deal of misuse and misinterpretation of these data. It is essential that the user understand how these data fit into the total scheme of fertility analysis in order to make maximum use of them.

An important warning must be made concerning the handling of the response, "no children ever born" for childless women. A special arrangement must be made to get this fact recorded. (Often a square or circle is provided that is to be checked if the respondent is

childless.) Otherwise when enumerating childless women there is a tendency just to leave the space blank, with the result that in data processing it is treated as "no information." Such understatements of childlessness distort the data seriously, especially for younger ages. This problem has been well discussed by El-Badry.<sup>3</sup>

Children-ever-born data (frequently hereinafter abbreviated as CEB) may be tabulated in two ways: (a) as distributions showing how many women have borne 0, 1, 2, 3, 4, 5, etc., children or (b) as cumulative totals of the number of children born. Table 18-18 is an illustration of the raw data that may be tabulated in this way, cross-tabulated by age:

- (a) they refer to fertility of real cohorts of women; hence they are longitudinal;
- (b) they refer to cumulative fertility, or fertility since first exposure to sex up to the time of enumeration;
- (c) they are retrospective and pertain to fertility over a varying (up to 30 year) span and have a vague time referent.

These, unfortunately, are almost the exact opposite specifications that are characteristic of vital statistics which lead to current fertility measures for hypothetical cohorts. It is for this reason that the demographers of the world should cease to place such exclusive emphasis upon children-ever-born statistics and permit "own children" statistics (discussed in the next section) to share the limelight. ("Own children" fertility statistics refer to the current fertility of hypothetical cohorts and are the direct census counterpart of vital statistics.) Misuse of CEB data arises most frequently from their being treated as if they refer to current fertility. On the other hand, much valuable information in these data is overlooked or neglected simply because they are not treated as longitudinal observations.

18-12a. Children-Ever-Born Rates. The basic children-ever-born rate is the number of children ever born per 1000 women. This may refer to all women aged 15-44 or to a particular age group.

(1) CEB Age Cumulative Fertility Rate is the number of children ever born per 1000 women of a certain age.

$$ACFR(CEB) = \frac{E}{F_x} \cdot \frac{1}{C} \cdot 1000 \quad (26)$$

where E is the cumulative number of children ever born,  
 C is the proportion of CEB that are enumerated by the census, and  
 F<sub>x</sub> is the number of women of age x.

Five year age groupings are usually employed. Two denominators can be used to compute these rates: (a) all women; and (b) all ever-married women (married, widowed, or divorced). The first yields general

Table 18 - 18. ILLUSTRATION OF PROCEDURE FOR CALCULATING CHILDREN-EVER-BORN RATES

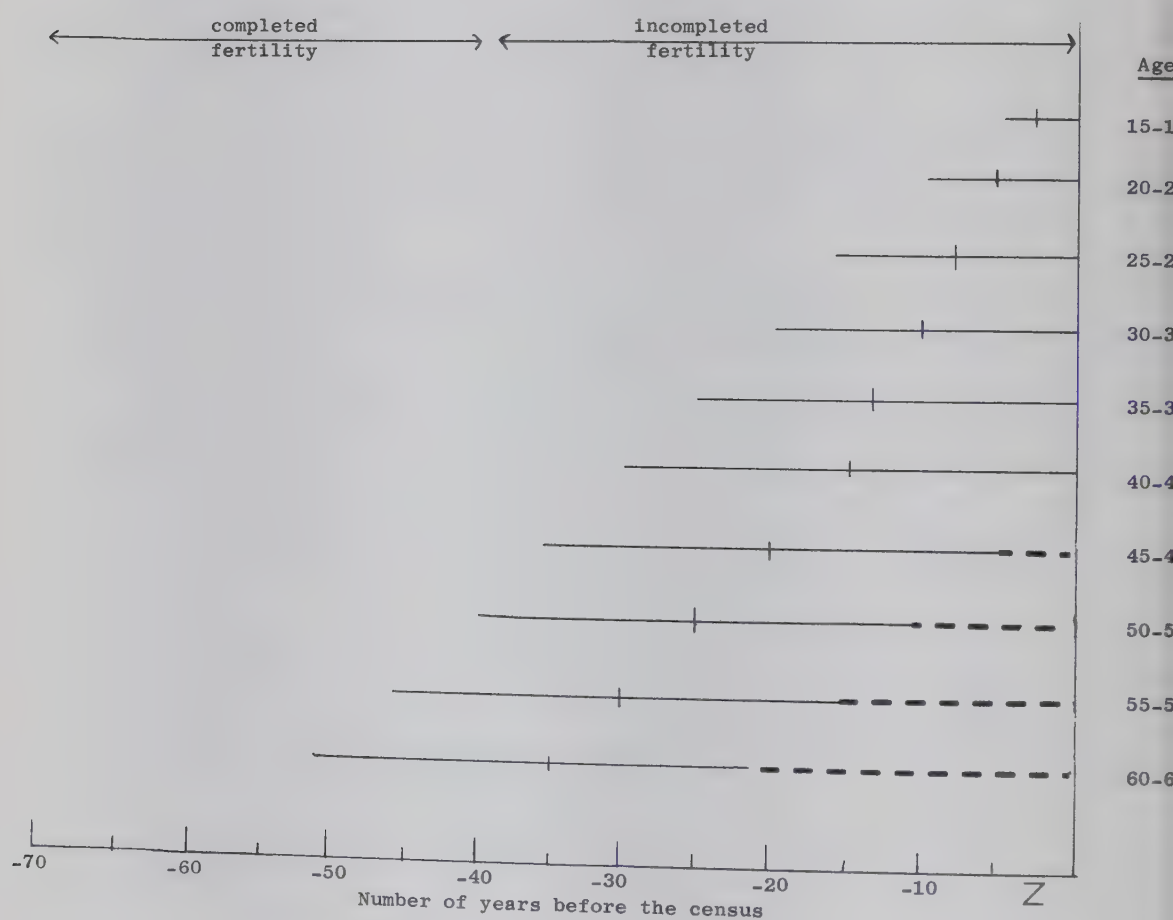
Age	Women (000)			Children ever born (000)		CEB per 1000 women		
	All women	Ever-married women	Women married husband present	To all ever married women	To women married husband present	All women	Ever-married women	Women married husband present
Total.....	63,563	52,572	35,020	131,752	85,529	2,073	2,506	2,442
15-44 years.....	36,069	27,219	21,628	62,991	50,288	1,746	2,314	2,325
15-19 years.....	6,588	1,061	856	839	657	127	791	767
20-24 years.....	5,506	3,934	3,344	5,671	4,758	1,030	1,441	1,423
25-29 years.....	5,537	4,956	4,127	11,116	9,228	2,007	2,243	2,236
30-34 years.....	6,100	5,681	4,553	14,957	12,111	2,452	2,633	2,660
35-39 years.....	6,439	6,049	4,650	16,212	12,725	2,518	2,680	2,737
40-44 years.....	5,898	5,538	4,098	14,197	10,808	2,407	2,564	2,637
45-49 years.....	5,560	5,200	3,653	12,491	9,031	2,247	2,402	2,472
50-54 years.....	4,927	4,553	2,984	10,737	7,240	2,179	2,359	2,426
55-59 years.....	4,406	4,043	2,411	10,088	6,117	2,290	2,495	2,538
60-65 years.....	3,719	3,432	1,794	9,310	4,921	2,503	2,713	2,742
65 years and over..	8,882	8,126	2,551	26,135	7,932	2,942	3,216	3,110

Note: Rate for all women is obtained by dividing children of all ever-married women by all women multiplied by 1,000.

Source: U.S. Bureau of the Census. U.S. Census of Population: 1960. Women by Number of Children Ever Born, Subject Report PC(2)-3A, 1964, Table 1.

rates while the second is an approximation of nuptiality rates. If additional specificity is desired, the denominator can be restricted to women who were married at the time of the census, or even further to women who were married and living with their husbands. In these last two cases, however, it is necessary to restrict the births in the numerators strictly to match the denominator, excluding children of widowed and divorced women and separated women. Table 18-18 illustrates the calculation of these rates.

Figure 18-4. TIME SPANS TO WHICH AGE TABULATIONS OF CHILDREN BORN STATISTICS REFER (Z is the date of the census)



Each five-year age group of these rates refers to a different time span. This fundamental and important feature is illustrated by the diagram of Figure 18-4. The average number of years preceding the census to which each age group refers is as follows:

Age	Years before the census	
	Average	Peak fertility
15-19.....	2.5	not yet reached
20-24.....	5.0	not yet reached
25-29.....	7.5	not yet reached
30-34.....	10.0	7.5
35-39.....	12.5	12.5
40-44.....	15.0	17.5
45-49.....	20.0	22.5
50-54.....	25.0	27.5
55-59.....	30.0	32.5
60-64.....	35.0	37.5

If we designate age 25 as the approximate peak childbearing (average) age of mothers at time of birth, we can then show the relationship of ASFR (CEB) to peak fertility. From this table we can see that at the time of the census, data for women of childbearing age (15-44) refer to a span of time 0-30 years in the past, each age referring to a different span. The average length of time is 15 years. For three of the age groups, peak childbearing has not yet occurred, while for the other age groups, it refers to years long past.

(2) CEB--Completed Fertility Rate. For ages 45 and over, the ACFR (CEB) is a measure of completed fertility because childbearing has ceased by the date of the census. By obtaining data for successively older cohorts of women one can presumably study trends or changes in completed fertility. However, completed fertility for the age group 40-44 refers to an average date of 15 years preceding the enumeration, and older cohorts refer to dates before that. Thus, recent changes in fertility could be expected to be reflected only in the data for cohorts with incomplete fertility. We use the symbols CFR (CEB) to identify completed fertility computed from children-ever-born data. There is a natural inclination for demographers to treat ACFR (CEB) and CFR (CEB) as if they were estimates of cross-sectional (current) ACFR and CFR. These assumptions are approximately justified only under conditions of unchanging fertility rates and complete enumeration of children ever born. Under conditions of changing fertility these assumptions are not justified.

One minor difference that mars the comparability of the CEB rates with their cross-sectional counterparts is the fertility of women who die before reaching a specified age. The fertility of these women is reflected in the vital statistics cross-sectional measures, but it is excluded from children-ever-born measures.

As will be discussed below, in many nations children-ever-born data are seriously under-reported. Many women fail to report infants born a long time ago--especially those who have died. There is also an apparent tendency to report only children from a current union and to omit illegitimate children and children of previous marriages. Accordingly, the value of C in equation (27) is almost certainly .95 or

smaller. In the absence of information about completeness of reporting, a correction factor of about this magnitude should be used.

(3) Age-Specific Fertility Rates from CEB Data. Mortara has suggested that children-ever-born data may be used to calculate age-specific fertility rates and total fertility rates simply by the process of successively subtracting the ACFR (CEB) rates for each younger age from the same rate for the next older age. This procedure is valid under conditions of unchanging fertility, such as have existed in the past in many developing countries. The procedure is expressed as:

$$\text{ASFR(CEB)} = A_x - A_{x+5} \quad (27)$$

where  $A_x$  equals ACFR(CEB) for age  $x$ , and

$A_{x+5}$  equals ACFR(CEB) for age  $x+5$ .

However, where there has been a recent change in fertility, this procedure is invalid. The fertility of the cohort of women aged 40-44, by definition, is the TFR obtained by this procedure, and the ASFR must by definition add up to this total. Thus, the fertility of women who were at their peak childbearing more than 15 years before the date of the census controls the results that will be obtained.

A refinement of the Mortara procedure, however, is to use the CEB data for two successive censuses. A procedure for establishing ASFR by this method (not reported elsewhere to the author's knowledge) is described below and illustrated in Table 18-19. Several steps are involved:

1. Record the CEB rates by single years of age for two successive censuses. If available only for five year periods reduce to single years by interpolation (columns 1 and 2).
2. Average the CEB rates for each age to obtain an average CEB rate that represents the midpoint of the intercensal period (column 3).
3. Compute the annual change in CEB rate during the intercensal period. This is the CEB rate at the end of intercensal period minus intercensal rate at the beginning, divided by 10 for each age (column 4).
4. Take one-half the intercensal change (Step 3) and subtract it from the average CEB rate (Step 2), to obtain an estimate of CEB rate 1/2 year prior to the midpoint of the intercensal period (column 5).
5. Add the annual change in CEB rate (Step 3) to the results of Step 4. This yields an estimate of CEB rate 1/2 year after the midpoint of the intercensal period (column 6).
6. From the CEB rate for age  $x$ , derived in Step 5, subtract

- the CEB rate for age  $x-1$  developed in Step 4. This is an estimate of the number of children each age group of women bore as they moved from average age  $x-1$  to age  $x$  (column 7).
7. Because the single year of age data refer to women whose average age is  $x$ , the results of Step 6 yield rates for age  $x+0.5$  years. To convert to average age  $x$ , it is necessary to average the rates for age  $x$  and  $x-1$ . For age 16, average the value of age 15 obtained in Step 2 and the value obtained for age 16 in Step 6, e.g.  $1/2 (16.3 + 10.5) = 13.4$  (the results are recorded in column 8).
  8. Step 7 will give an estimated value of ASFR. However, because of incompleteness of enumeration of CEB, the rates will tend to be too low. We must, therefore, inflate the values by a factor that raises them to the proper average level. This may be done by applying the ASFR estimated in Step 7 to the age distribution of women at the census to estimate an average GFR, based on ASFR from CEB data. The ratio of the "true" GFR (estimated by procedures outlined in 18-10) to this estimated GFR constitutes a correction factor which is then applied to each ASFR individually. Alternatively, a "true" value of TFR, estimated by other procedures, may be used as the numerator of an inflation factor in which TFR implied by the ASFR is the denominator.
  9. Convert the ASFR for single years into five year age groupings. This may be done by adding the rates into five year age groups and dividing by five.

As Table 18-19 indicates, this procedure yields results that are surprisingly close to the actual average ASFR for 1950-60. ASFR for girls 15-19 years of age is seriously underestimated, while there is a tendency to overestimate ASFR for ages 30 and above. This was a decade of substantial change in fertility patterns. If the Mortara (single census) procedure is used either for 1950 or 1960, results are obtained that are nonsensical.

General CEB Rates and their Standardization. Frequently in fertility studies a "general rate of children-ever-born" is encountered. This is the number of children-ever-born per 1000 women of childbearing age.

$$\text{GFR(CEB)} = \frac{\sum_{15}^{44} \text{CEB}}{C_{30}F_{15}} \cdot 1000 \quad (28)$$

This is a "mongrel" rate, which has no counterpart in the system of fertility measurement introduced in Section 18-3. It states the average number of children women of childbearing age have borne.

Table 18-19. PROCEDURE FOR ESTIMATING AGE-SPECIFIC FERTILITY RATES FROM DATA ON CHILDREN-EVER-BORN, USING CEB RATES FROM TWO SUCCESSIVE CENSUSES; UNITED STATES POPULATION: 1950 AND 1960

Age	Children-ever-born per 1000 total women of age		Mid-point CEB rate	Annual inter-censal change in CEB rate	CEB RATE FOR MID-CENSAL YEAR		ESTIMATED ASFR-- PRELIMINARY		Final estimate of ASFR*
	1950	1960			one-half year prior mid-point	one-half year after mid-point	For age X+0.5	For age X	
15 years.....	5	16	10.5	1.1	9.9	11.0	10.5	5.6	6.0
16 years.....	19	32	25.5	1.3	24.9	26.2	16.3	13.4	14.4
17 years.....	55	77	66.0	2.2	64.9	67.1	42.2	29.3	31.5
18 years.....	124	166	145.0	4.2	142.9	147.1	82.2	62.2	66.8
19 years.....	240	330	285.0	9.0	280.5	289.5	146.6	114.4	122.8
20 years.....	384	543	463.5	15.9	455.6	471.5	191.0	168.8	181.2
21 years.....	538	760	649.0	22.0	637.9	660.1	204.5	197.7	212.2
22 years.....	693	997	845.0	30.4	829.8	860.2	222.3	213.4	229.1
23 years.....	858	1239	1048.5	38.1	1029.5	1067.6	237.8	230.1	247.0
24 years.....	1008	1455	1231.5	44.7	1209.2	1253.9	224.4	231.1	248.1
25 years.....	1173	1659	1416.0	48.6	1391.7	1440.3	231.1	227.5	244.2
26 years.....	1287	1839	1563.0	55.2	1535.4	1590.6	198.9	215.0	230.8
27 years.....	1423	1992	1707.5	56.9	1679.0	1735.9	200.5	199.7	214.4
28 years.....	1530	2106	1818.0	57.6	1789.2	1846.8	167.8	184.2	197.7
29 years.....	1638	2188	1913.0	55.0	1885.5	1940.5	151.3	159.6	171.3
30 years.....	1687	2310	1998.5	62.3	1967.0	2029.3	143.8	147.6	158.4
31 years.....	1801	2350	2075.5	54.9	2048.0	2102.9	135.9	139.8	150.1
32 years.....	1857	2422	2139.5	56.5	2111.3	2167.8	119.8	127.8	137.2
33 years.....	1941	2450	2195.5	50.9	2170.1	2221.0	109.7	114.8	123.2
34 years.....	1967	2449	2208.0	47.2	2184.4	2231.6	61.5	85.6	91.9
35 years.....	1942	2458	2200.0	51.6	2174.2	2225.8	41.4	51.4	55.2
36 years.....	2022	2480	2251.0	45.8	2228.1	2273.9	99.7	70.6	75.8
37 years.....	2070	2477	2273.5	40.7	2253.2	2293.9	65.8	82.8	88.9
38 years.....	2023	2495	2259.0	47.2	2235.4	2282.6	29.4	47.6	51.1
39 years.....	2102	2444	2273.0	34.2	2255.9	2290.1	54.7	42.1	45.2
40 years.....	2056	2423	2239.5	36.7	2221.2	2257.9	2.0	28.4	30.5
41 years.....	2130	2403	2266.5	27.3	2252.9	2280.2	59.0	30.5	32.7
42 years.....	2157	2360	2258.5	20.3	2248.4	2268.7	15.8	37.4	40.1
43 years.....	2160	2340	2250.0	18.0	2241.0	2259.0	10.6	13.2	14.2
44 years.....	2177	2279	2228.0	10.2	2222.9	2233.1	- 7.9	1.3	1.4

\*An inflation factor of 1.0735 was used to bring estimated ASFR to average level of "true" fertility 1950-60.

Summary	Estimated 1950-60	Official 1950-60	Difference
15-19 years....	48.3	90.6	- 42.3
20-24 years....	223.5	236.1	- 12.6
25-29 years....	211.7	187.9	23.8
30-34 years....	132.2	113.9	18.3
35-39 years....	63.4	57.2	6.2
40-44 years....	23.8	16.9	6.9

Source: Data for 1950 from U.S. Census of Population: 1950, Special Reports, Fertility, Table 8, page 5C-28. Data for 1960 from U.S. Census of Population: 1960, Subject Reports, Women by Number of Children Ever Born, Table 4, page 10

It is not a probability measure and, in fact, has no significant theoretical meaning. Its use should be avoided because it is extremely sensitive to age composition: a young age composition yields a low rate while an older age composition yields a higher rate. This is the exact reverse of the age-differential for the basic rates and creates serious problems in analysis. Inasmuch as younger populations tend to have lower current fertility than older populations when fertility is declining, this rate can show results that are the exact reverse of those that would be shown by valid current fertility measures. GFR(CEB) should be used only when standardized for age composition.

And this rate is not easy to standardize. Because the younger ages have such small numbers in relationship to the older ages, a simple straightforward application of these rates to a standard population would yield a result that is both meaningless and determined largely by the relative proportion of women at older ages. An appropriate procedure is to compute quasi-age-specific fertility rates (as described in equation (27) and more precisely established in Table 18-19), to apply these to a standard population of women aged 15-44 years, and then to cumulate. For indirect standardization, a set of quasi-ASFR derived by the procedure of Table 18-19 are applied to the age composition of a particular population.

18-13. Calculation of Current Birth Rates from "Own Children" Tabulations of Census Data. A powerful technique for measuring current fertility is the "own children" procedure first proposed by Wilson Grabill and later refined jointly by Grabill and Lee-Jay Cho.<sup>4</sup> By a very simple recoding of information contained on census schedules, this procedure yields data which can be used to derive highly precise measures of age-specific fertility, GFR, TFR, and CBR. In addition, these measures may be computed for any characteristic of the population (education, income, occupation, etc., of either the mother or the father) or any combination of these characteristics. It is, therefore, the most direct and reliable method for obtaining data on current differentials in fertility.

1. Children under 5 years of age enumerated at a census represent the survivors of all live births born during the five years preceding the census. Children aged 5-9 years at a census represent the survivors of all live births born during the period 5 to 10 years preceding the census.
2. All but a small fraction of children under 10 years of age reside in the same household as their mother and father, and they will be enumerated simultaneously with their parents at the time of the census.
3. Census enumeration is done by households, with every person enumerated being identified in terms of his rela-

tionship to the household head. With this information it is possible to specify, for every child, which adult is the mother and which is the father, if the parent(s) are present in the household.

4. Therefore, census enumeration schedules may be looked upon as a set of birth registers, in which live births have been recorded together with the characteristics of both parents.
5. These registers are deficient in minor respects, but for each deficiency an appropriate correction or adjustment can be made.
  - (a) Children who have died can be estimated by use of life tables.
  - (b) The small proportion of children who live away from their parents can be determined and an adjustment can be made.
  - (c) The tendency to undercount young children at the time of the census can be corrected by an adjustment.

The procedure calls for no additional questions on the census schedules. After the census has been taken, it demands only a special coding, either for all or for a sample of households. It may be performed on data from ancient censuses or sample surveys, provided that the original schedules are available. This procedure places in the hands of most nations an opportunity to measure, with at least moderate reliability, its level of current fertility (and differentials in fertility) once at each census enumeration. These measures are independent of the quality of the nation's vital registration system. It would be difficult to overemphasize, therefore, the importance of making this procedure a routine part of every census operation.

18-13a. Relationship Between Age and Time. In order to avoid confusion in the formulas developed below, it is important to understand the precise nature of the correspondence between census age data and the span of time to which the fertility estimates refer. We assume that all age data are collected in terms of "age last birthday." We identify by the symbol  $Z$  the span of time less than 12 months preceding the census enumeration, by  $Z-1$  the span of time 12 months but less than 24 months preceding the census enumeration, etc. Time in this case does not refer to calendar years, but to intervals of 12 months preceding the census. In the formulas which follow, we will use the symbol "a" to refer to the age of children, and the symbol "x" to refer to the age of mothers. There is an exact correspondence between age of the child and the year to which the estimates refer. This is given by the relation:

$$\text{Years to which the estimates refer} = Z - a$$

The following diagram should clarify the situation fully.

TIME	Year to which estimate refers									
	Z	Z-1	Z-2	Z-3	Z-4	Z-5	Z-6	Z-7	Z-8	Z-9
AGE	0	1	2	3	4	5	6	7	8	9
Age of child at last birthday										

From this diagram it may readily be seen that the data for "own children under 5" will yield an average fertility rate for a span covering five years preceding the census, while data on "own children aged 5-9" may be used to obtain data for fertility 5 to 10 years preceding the census. It may further be seen that if data on "own children" are collected for single years of age they may be converted into estimates of fertility for each calendar year for the entire decade preceding the census. Thus, own children data may be used not only to establish the level of current fertility, but also to measure recent changes or trends in fertility.

18-13b. Coding Procedure. A probability sample of households will usually be drawn from the complete census returns because the cost of coding and processing can be greatly reduced by sampling. For this sample the following steps are performed.

1. Set up a tabulation card for each woman aged 15 to 49 in the sample. This tabulation card will have three sections:

Own children register

Characteristics of the woman

Characteristics of the husband (if residing in the same household)

(a) The "own children register" section should consist of 12 fields as follows:

<u>Field number</u>	<u>Number of own living children whose age last birthday was:</u>
1.....	Under 1 year
2.....	1 year
3.....	2 years
4.....	3 years
5.....	4 years
6.....	5 years
7.....	6 years
8.....	7 years
9.....	8 years
10.....	9 years
11.....	Under 5 years (summary)
12.....	5 to 9 years (summary)

A 13th field, specifying total number of live-born children ever born (irrespective of whether or not they are living or their present

age or place of residence at time of census) should be coded also. A code of zero in this field means that the woman has remained childless up until the time of the census.

(b) The "characteristics of the woman" section consists of a series of fields, each devoted to reporting a characteristic of the woman. The fields established must depend upon the items of information for which data were collected by the census. Possible characteristics are:

Age at time of census	Occupation
Race or ethnic origin	Income in previous year
Marital status	Place of birth (state, urban-rural)
Age at first marriage	Migration status
Number of times married	Religion
Employment status	

(c) The "characteristics of the husband" section consists of a series of fields, each devoted to reporting a characteristic of the husband of the woman. The items selected must depend upon the data that were collected at the census. They may be identical to those listed above for the woman.

2. Establish rules for identifying the mothers of children under 9 years of age. The coding is done one household at a time. For each household the coder identifies every woman aged 15 to 49 and then searches the schedule to learn how many children under 9 years of age she has born. In each case an effort is made to "match" mother and child from the interview schedule, using information of "relationship to head" and age-sex data. Typical matches are:

<u>Relationship</u>	<u>Age</u>	<u>Sex</u>
Wife of head	15-54	Female
Child of head	0- 9	Male or Female
Child of head	15-54	Female
Grandchild of head	0- 9	Male or Female
Daughter-in-Law	15-54	Female
Grandchild of head	0- 9	Male or Female
Other relative	15-54	Female
Other relative	0- 9	Male or Female
Lodger	15-54	Female
Lodger	0- 9	Male or Female

The last two pairs listed above are ambiguous; they may or may not be a mother-child pair. Additional information may be obtained from the schedule that will assist in determining whether they are or are

not such a pair. The information for marital status, place of birth, and ethnic origin can sometimes help. Most households will be similar to the first or second of the above examples and will present few problems. As special situations are encountered, definitions that are nonbiasing must be generated.

A check on the matching should be made to assure that a mother has been assigned to every child if possible and that no child has been assigned to more than one mother.

3. Assign each mother-child pair to its proper age-year field. In fields of 1-10, codes of 0, 1 or 2 will be the most frequent entries. (Code 2 is possible in the case of birth of twins or two births in one year.) In fields 11 and 12, the codes 0, 1, 2, 3, 4 will occur most commonly.

4. For each woman, assign a code of 0 in fields not already identified as a mother-child pair. Check to make sure that the sum of fields 1-5 is equal to field 11, and that the sum of fields 6-9 is equal to field 12.

5. Transfer to the woman's card previously coded information concerning her characteristics and her husband's characteristics.

18-13c. Tabulation Procedure. It will become clear below that the data needed to compute fertility rates from own children data are:

- (a) Number of own children (numerator of the rates),
- (b) Number of women (denominator of the rates),
- (c) Number of own children per 1000 women.

This information may be tabulated for each of the 13 fields outlined above or only for field 11 or 12. The more detailed the information available, the more detailed the estimates can be.

The numerators (number of own children) are obtained by cumulating the fields 1-12. The denominators (number of women) are obtained by counting the number of women of childbearing age in the sample. These counts may be of several types:

- (a) Total, all women;
- (b) Age-specific, separately for each age group of women;
- (c) Trait-specific, separately for each trait group of women or their husbands;
- (d) Age-trait specific, separately for each age group further subdivided by trait groups.

A computer program, TASA, is being published as Manual Number 9 in the series Family Planning Research and Evaluation Manuals. This program performs the numerator-denominator-rate tabulation for any combination of age-specific and trait-specific classifications desired. If the researcher does not already possess a computer program capable of doing this, he will find it useful to use the TASA program.

18-13d. Adjustment to Correct for Deficiencies. Although it will be shown below that adjustment for all known deficiencies can be performed as a single step, in order to clarify the procedure, the adjustment for each type of deficiency will be shown separately.

(1) Adjustment for Children Who Do Not Live with Their Mothers. The sum of all own children of each age (sum for the entire nation or other area) is expressed as a ratio of all children of the same age reported at the same census.

Let  $P'_a$  be the number of own children of any age who are tabulated as living with their parents, and  
 $P_a$  be the total number of children of the same age, counted at the same census

$P'_a$  is obtained by summing the data for all women in the sample; it is the numerator of the own child/woman ratio.  $P_a$  is obtained from official census tabulations; it is the total child population of the specified age.

$$R_a = \frac{P_a}{P'_a} \quad (29)$$

where  $R_a$  is the adjustment for children who do not live with their mothers.

$R_a P'_a$  is an unbiased estimate of all children of age  $a$  counted at the census.

For most nations of the world it will be found that 95 percent or more of children under 5 live with their mothers. Among Negro and certain other populations of the United States, the Caribbean, and perhaps Africa, it will be found that as much as 15 percent or more live apart from their parents. Under conditions of high proportions of children living away from home, the adjustment factor for younger women will be much greater than that for older women. Hence,  $R_a$  needs to be computed on an age-of-mother basis. A procedure for doing this is described in Cho, Grabill, and Bogue.<sup>5</sup>

(2) Adjustment for Women of Childbearing Age Not Living in Households. Just as some children may not live with their mothers, not all women of childbearing age will be found living in private households, or the sampling of women chosen for measurement of fertility may not correspond exactly to the exact count of women of childbearing age.

Let  $F'_x$  be the number of women of any age who are tabulated or estimated from the special coding for "own children"; and  
 $F_x$  be the total number of women of the same age counted at the census.

$$R_x = \frac{F_x}{F'_x} \quad (30)$$

where  $R_x$  is the adjustment for women who do not live in households.

$R_x F'_x$  is an unbiased estimate of women of childbearing age counted at the census.

(3) Adjustment for Undercounting of Young Children and of Women. The net undercount made on almost all censuses both of young children and women of childbearing age affects the estimates of fertility based on own children. If both children under age 5 and women of associated ages are missed in equal proportions, ratios of young children to women would not be affected by the adjustments. Usually, if an undercount occurs, it will be found that children suffer from a larger undercount than women.

Let  $C_a$  be an estimate of the proportion of children aged  $a$  years enumerated at the census, and  $C_x$  be an estimate of the proportion of women aged  $x$  enumerated at the census then an unbiased estimate of the correct number of children.

$$P_a^* = \frac{P_a' R_a}{C_a} \quad (31)$$

where  $P_a^*$  is an unbiased estimate of the true number of children residing in the territory at the time of the census.

$$F_x^* = \frac{F_x' R_x}{C_x} \quad (32)$$

where  $F_x^*$  is an unbiased estimate of the true number of women of childbearing age  $x$  residing in the territory at the time of the census.

(4) Adjustments for Mortality of Children. In order to convert the estimates of living children into estimates of births that occurred during the years immediately preceding the census, it is necessary to "reverse survive" the data for children by means of life table survival ratios.

Let  ${}_a S_0$  be the probability of surviving from birth to be in year of age  $a$  at the time of the census count.

If children under  $a$  years of age are being used to estimate fertility then the appropriate formula for calculating the survival ratio is

$${}_a S_0 = \frac{\sum_{0}^{a-1} L_a}{a l_0} \quad (33)$$

where  $l_0$  is the radix of the life table (usually 100,000), and  $\sum_{0}^{a-1} L_a$  is the life table stationary population under age  $a$ .

Specifically, if a survival ratio for children under 5 is computed:

$${}_5S_0 = \frac{L_0 + L_1 + L_2 + L_3 + L_4}{500,000} = \sum_0^4 L_x / {}_5l_0 \quad (34)$$

$$\text{an equivalent value of which is } \frac{T_0 - T_5}{{}_5l_0} \quad (34a)$$

An unbiased estimate of the number of births that occurred in the ath year before the census count is

$$B^* = \frac{P_a^*}{{}_aS_0} = P_a^* \frac{\sum_0^4 L_x}{{}_5l_0} \quad (35)$$

(5) Adjustments for Mortality of Women. The women who are enumerated in a census are the survivors of the larger group of women who bore the "own children" estimated by this procedure. It is necessary to convert the count of surviving women into an estimate of the childbearing population. This is done by a procedure of reverse survival comparable to that described above for children.

Let  ${}_aS_{x-a}$  be the probability of surviving from birth of a child aged a to be in year of age x at the time of the census count =  $L_x / L_{x-a}$

$$\text{Then } F = \frac{F_x^*}{{}_a/2 S_{x-a}} \quad \text{is an unbiased estimate of the number of women of childbearing age } x \text{ during the span of time to which the estimates refer.} \quad (36)$$

If children under a years of age are being used to estimate fertility, then one-half of this span is the average amount of time which women have survived since the birth of the children, or a/2. For the youngest ages and the oldest ages this assumption is incorrect. At the youngest ages, average survival time is less than a/2; and at the oldest ages it is greater than a/2. However, these biases have an insignificant impact upon the overall result.

The procedures described above are illustrated in Table 18-20. In column 2 of this worksheet are recorded the number of own children under 5, by age of woman, taken from fertility tabulations of the 1950 census. In the first column of this worksheet are reported the numbers of women of each age, as officially reported at the census. In columns 3 and 4 the count of children and women have been adjusted for known errors, as explained in the notes at the foot of the table. These counts are then inflated by survival ratios to take account of mortality, both among children and women. Survival factors used represent mortality over a 2.5 year period, since this is the average interval between birth and date of census. The last column in the lower panel of the table represents the fully adjusted data.

18-13e. Conversion of Ratios of Children to Women into Estimates of ASFR. The ratios calculated in Table 18-20 represent

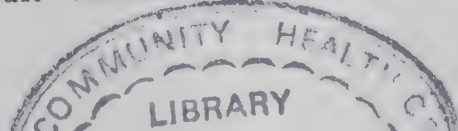
Table 18-20. WORKSHEET FOR COMPUTING CORRECTED RATIOS OF OWN CHILDREN TO WOMEN FROM CENSUS DATA; UNITED STATES: 1950

Age of woman at census	Census data unadjusted (000)		Census data adjusted for net undercount		Survival ratios	
	Women	Own children	Women	Own children	Women	Own children
Total.....	38,764	15,623	39,539	17,096	....	....
15-19 years.....	5,322	495	5,428	542	.9976	.9674
20-24 years.....	5,878	3,586	5,996	3,925	.9965	.9674
25-29 years.....	6,277	4,948	6,403	5,415	.9961	.9674
30-34 years.....	5,897	3,498	6,015	3,828	.9953	.9674
35-39 years.....	5,712	2,077	5,826	2,273	.9935	.9674
40-44 years.....	5,125	847	5,228	927	.9901	.9674
45-49 years.....	4,553	170	4,644	186	.9846	.9674
<b>Adjusted Data</b>						
	Women		Own children		Own children per 1000 women	
Total.....	39,790		17,671		....	
15-19 years.....	5,441		560		103.1	
20-24 years.....	6,017		4,057		674.2	
25-29 years.....	6,428		5,597		870.7	
30-34 years.....	6,043		3,957		654.8	
35-39 years.....	5,864		2,350		400.8	
40-44 years.....	5,280		958		181.4	
45-49 years.....	4,717		192		40.7	

Source: Special Reports: Fertility, Table 34, page 5C-154. Data on age are official age distributions reported in U.S. Summary: Characteristics of the Population

Notes: (1) Women are assumed to have suffered a net undercount of 2 percent. Children are assumed to have suffered a net undercount of 5 percent. Own children under 5, totalling 15,623,000, were inflated by a factor of 1.0397 to bring them to the census count of all children under 5. Hence, column 3 is column 1 multiplied by 1.02, and column 4 is column 2 inflated by a combined factor of 1.0944. The data in columns 1 and 2 of the lower panel are the figures in columns 3 and 4 of the upper panel divided by the respective survival ratios of columns 5 and 6.

the cumulative fertility of five years, during which the population has aged. The task of disentangling this cumulation into an average schedule of ASFR for one year with correct assignment of ages to conventional 5-year age intervals is a somewhat complex one. Fortunately, Grabill and Cho have simplified it by establishing a set of coefficients which accomplish several tasks in one step. These coefficients are reported in Table 18-21. The procedure by which they are used is described at the foot of that table. The use of these coefficients to estimate a schedule of ASFR from the ratios computed in Table 18-20 is illustrated in Table 18-22. If the reader wishes a full exposition of the adjustments that are accomplished by the coefficients of Table 18-21 (which are Sprague fifth difference osculatory interpolation multipliers) he should consult "Methodology for the Measure-



ment of Current Fertility from Population Data on Young Children," by Wilson H. Grabill and Lee-Jay Cho, Demography, vol. 2, 1965, pp. 50-73.

Table 18-21. COEFFICIENTS FOR ESTIMATING AVERAGE ANNUAL AGE-SPECIFIC FERTILITY RATES FROM CENSUS-BASED RATIOS OF CHILDREN UNDER 5 YEARS OF AGE PER 1,000 WOMEN, BY AGE OF WOMAN AT END OF THE PERIOD

Fertility rate to be estimated	Ratios of children 0-4 years to women, by age of woman						
	15-19	20-24	25-29	30-34	35-39	40-44	45-49
	ratio	ratio	ratio	ratio	ratio	ratio	ratio
ASFR 15-19 years.....	.1378	.1109	-.0208	.0020	.0002	-.0001	.....
ASFR 20-24 years.....	-.0365	.1378	.1109	-.0208	.0020	.0002	-.0001
ASFR 25-29 years.....	.0073	-.0365	.1378	.1109	-.0208	.0020	.0002
ASFR 30-34 years.....	-.0008	.0073	-.0365	.1378	.1109	-.0208	.0020
ASFR 35-39 years.....	.0001	-.0008	.0073	-.0365	.1378	.1109	-.0208
ASFR 40-44 years.....	.....	.0001	-.0008	.0073	-.0365	.1378	.1109
ASFR 45-49 years.....	.....	.....	.0001	-.0008	.0073	-.0365	.1378

The ratios to be inserted in the first line of this table are age-specific ratios of children under 5 to women of the 5-year age group indicated by the column heading, adjusted for mortality of children and women since birthdate of children, for proportion of population under 5 not living with mother, and for undercount of women and children in the census, so that results are total births per 1,000 women in 5-year period preceding the census by age of woman at the end of the period.

The coefficients are used by inserting a schedule of ratios in the first line of the table, multiplying the ratio in each column by the coefficients in that column, recording each result in its appropriate cell. The products are then summed by rows, taking account of signs. The sum for each row is an estimated ASFR for the age group indicated.

The lower panel of Table 18-22 compares the estimated ASFR with the five-year average of officially reported ASFR (corrected for underregistration). The amazingly close match demonstrates the great power of this technique. Moreover, this example was deliberately chosen because it covers one of the most dramatic fertility changes in recent demographic history anywhere in the world: the onset of the famed "baby boom" in the United States. In 1946-47 the birth rates of the United States rose dramatically over the levels of 1945. They continued to rise through 1950. The average ASFR for the period therefore represents a very unstable condition. Nevertheless, the "own children" procedure was able to measure this average with almost zero overall error in TFR, and with very small errors at the individual age groups. Only for the ages above 40 (where birth rates are low) are the relative errors substantial, and the absolute errors are small for all age groups.

With the schedule of ASFR established, TFR may be calculated directly by summing ASFR and multiplying by 5. GFR may be obtained by applying the schedule of ASFR to the count of women of childbearing age to get expected births, then dividing by all women 15-44 or 15-49. CBR may be obtained by dividing expected births

Table 18-22. ILLUSTRATION OF USE OF COEFFICIENTS OF TABLE 18-21 TO DERIVE AGE-SPECIFIC FERTILITY RATES FROM RATIOS OF CHILDREN TO WOMEN DERIVED IN TABLE 18-20 FROM CENSUS TABULATIONS OF OWN CHILDREN

Estimate for:	Ratios of children 0-4 years of age to women, by age						Estimated ASFR (sum of rows)
	15-19 years	20-24 years	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years
Total.....	103.1	674.2	870.7	654.8	400.8	181.4	40.7
ASFR 15-19 years.....	14.21	74.77	- 18.11	1.31	.08	- .02	....
ASFR 20-24 years.....	3.76	92.90	96.56	-13.62	.80	.04	.00
ASFR 25-29 years.....	.75	-24.61	119.98	72.62	- 8.34	.36	.01
ASFR 30-34 years.....	- .08	4.92	- 31.78	90.23	44.45	- 3.77	.08
ASFR 35-39 years.....	.01	- .54	6.36	-23.90	55.23	20.12	- .85
ASFR 40-44 years.....	....	.07	- .70	4.78	-14.63	25.00	4.51
ASFR 45-49 years.....	....	....	.09	- .52	2.93	- 6.62	5.61
							72.24
							180.44
							160.77
							104.05
							56.42
							19.03
							1.49

Comparison of estimated ASFR with official ASFR

Age	Official ASFR	Estimated ASFR	Difference	
			Absolute	Relative
15-19 years.....	71.8	72.2	+0.4	+.006
20-24 years.....	186.2	180.4	-5.8	-.031
25-29 years.....	159.6	160.8	+1.2	+.007
30-34 years.....	105.4	104.1	-1.3	-.012
35-39 years.....	56.5	56.4	-0.1	-.002
40-44 years.....	16.1	19.0	+2.9	+.180
45-49 years.....	1.4	1.5	+0.1	+.071

by the total population.

To repeat, the above procedure may be carried out for any trait-group desired--occupation, education, income, area-of-residence, etc. It offers, therefore, a wonderful avenue for computing conventional measures of current fertility for highly specific subcategories of the population. Its utility in the study of differential fertility is therefore obvious. Why this procedure, so inexpensive, so precise, and so powerful, remains comparatively unused throughout the world is one of the mysteries of modern demography.

18-14. The Measurement and Explanation of Differential Fertility. Within every population there tend to be substantial differences in fertility; some subgroups bear children at significantly higher (or lower) rates than average. An important segment of fertility analysis is to detect these "fertility differentials" and to explain them.

An example of differential fertility is one that associates childbearing with race. Table 18-23 reports the basic fertility rates for the white and nonwhite population of the United States in 1960 and 1966. In terms of all of the measures it is evident that nonwhite women are much more fertile than are white women. These differences refer to particular years and, therefore, are cross-sectional.

Table 18-23. DIFFERENTIAL FERTILITY OF WHITE AND NONWHITE POPULATIONS OF THE UNITED STATES: 1960-66

Age and fertility measure	1960		1966		Difference, 1960		Change, 1960-66	
	White	Nonwhite	White	Nonwhite	Absolute	Percent	White	Nonwhite
<u>Age-specific rates:</u>								
15-19 years.....	79.4	158.2	60.8	135.5	74.7	122	-18.6	-22.7
20-24 years.....	252.8	294.2	179.9	228.9	49.0	27	-72.9	-65.3
25-29 years.....	194.9	214.6	146.6	169.3	22.7	15	-48.3	-45.3
30-34 years.....	109.6	135.6	82.7	107.9	25.2	30	-26.9	-27.7
35-39 years.....	54.0	74.2	40.0	57.7	17.7	44	-14.0	-16.5
40-44 years.....	14.7	22.0	10.8	18.4	7.6	70	- 3.9	- 3.6
45-49 years.....	0.8	1.7	0.7	1.4	0.7	100	- 0.1	- 0.3
General fertility rate.....	113.2	153.6	86.4	125.9	39.5	46	-26.8	-27.7
Total fertility rate.....	3533	4522	2609	3614	1005	39	-924	-908

Source: Vital Statistics of the United States: 1966

Just as for the measurement of fertility change, the analysis of fertility differentials may be done in terms of absolute differences or in terms of relative (percent) differences. Thus, we can say (using TFR as an example) that in 1966 the average size of the completed nonwhite family was 1 child greater than the average size of the white completed family. This states the differential in absolute terms. Alternatively, we could say that the TFR for the nonwhite population was 39 percent greater than the TFR for the white population. This states the differential in relative terms. The relative differentials are es-

pecially useful for studying the age patterns of differentials. In this example, we can see that the relative differentials by race are greatest at the youngest and oldest ages.

The best overall measure of differential fertility is age-standardized GFR, standardized on an age composition intermediate between the two populations being compared. (National population as of the year of comparison is a good choice.) No overall fertility differential could be considered to be "refined" until it has been controlled for age, either by use of GFR standardized or TFR.

Another aspect of fertility differentials that needs to be explored is nuptiality. Is a fertility difference due to differences in the rate of childbearing within marriage, or simply to differences in the proportion married? Usually, demographers will wish to eliminate the effects of both age and nuptiality in studying fertility differentials; otherwise all other explanations are ambiguous. The solution, of course, is the use of ASFR(N) for each category of the differential and the comparison of GFR(N) standardized for age.

Equally as important as the measurement of fertility differentials is the measurement of change in differentials. Are the differences increasing or decreasing over time? Table 18-23 presents data for two dates in order to illustrate the solution to this problem. When studying change in differentials, great emphasis needs to be placed on the absolute differentials, rather than upon the relative differentials. For example, between 1960 and 1966 the relative differential between the white and nonwhite populations changed very little. Despite this fact, both of them underwent very substantial absolute declines. For most ages, the absolute declines for the nonwhite population were equal to or slightly greater than for the white population. This means that the growth rates of the two populations slackened by about the same amount. Statistics of relative change in differentials should be interpreted with caution, for they may mean comparatively little. For example, the fertility rates of the white and nonwhite populations could both decline to zero in such a way that up until the moment they both became zero the nonwhite would have a birth rate 50 percent greater than whites.

Use of Children-Ever-Born Data for the Study of Differential Fertility. Many forces that lead to differential fertility act in a sustained way, over many years. In such cases, their cumulative effect upon fertility should be much more readily observable than their cross-sectional effect at a particular date. For this reason, children-ever-born data are particularly useful for studying differential fertility. Table 18-24 illustrates the use of these data to study educational differentials in fertility. The upper panel of this table reports the CEB rates, as given in the census. In the center panel these rates

Table 18-24. USE OF CHILDREN-EVER-BORN DATA TO MEASURE FERTILITY DIFFERENTIAL; EDUCATIONAL ATTAINMENT OF THE UNITED STATES POPULATION: 1960

Educational attainment	Age of woman							
	15-19 years	20-24 years	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50+ years
Total.....	725	1370	2171	2564	2625	2515	2354	2753
No school years completed...	1368	2208	3236	4009	4148	4326	4121	4396
Elementary: 1 to 4 years...	1212	2101	2978	3650	3836	3856	3810	4136
5 to 7 years...	1067	1976	2765	3112	3259	3155	3014	3369
8 years.....	965	1891	2563	2829	2892	2769	2601	2855
High school: 1 to 3 years...	799	1779	2479	2679	2685	2539	2332	2387
4 years.....	470	1202	2064	2435	2450	2275	2013	1966
College: 1 to 3 years...	383	890	1872	2405	2450	2279	1970	1873
4 years.....	....	543	1445	2244	2445	2278	1900	1686
5 years or more	....	458	1039	1798	2072	1969	1652	1370
Ratio to average rate for age group								
No school years completed...	1.89	1.61	1.49	1.56	1.58	1.72	1.75	1.60
Elementary: 1 to 4 years...	1.67	1.53	1.37	1.42	1.46	1.53	1.61	1.50
5 to 7 years...	1.47	1.44	1.27	1.21	1.24	1.25	1.28	1.22
8 years.....	1.33	1.38	1.18	1.10	1.10	1.10	1.10	1.04
High school: 1 to 3 years...	1.10	1.30	1.14	1.04	1.02	1.01	0.99	0.87
4 years.....	0.65	0.88	0.95	0.95	0.93	0.90	0.86	0.71
College: 1 to 3 years...	0.53	0.65	0.86	0.94	0.93	0.91	0.84	0.68
4 years.....	....	0.39	0.67	0.88	0.93	0.91	0.81	0.61
5 years or more	....	0.33	0.48	0.70	0.79	0.78	0.70	0.50
Summary	Young (20-29 years)		Intermediate (30-44 years)			Older (over 45 years)		
No school years completed...	1.55		1.62			1.68		
Elementary: 1 to 4 years...	1.45		1.47			1.56		
5 to 7 years...	1.36		1.23			1.25		
8 years.....	1.28		1.10			1.07		
High school: 1 to 3 years...	1.22		1.02			.93		
4 years.....	.91		.93			.79		
College: 1 to 3 years...	.76		.93			.76		
4 years.....	.53		.91			.71		
5 years or more	.41		.76			.60		

have all been expressed as ratios of the average rate for the age group. Thus, a ratio above 1.0 indicates above average fertility and a ratio below 1.0 indicates below average fertility. The extent of deviation from 1.0 reveals the relative degree of difference. Thus, by expressing the data as ratios, all cells of the table become comparable to each other. The pattern of relative differentials becomes even clearer when summarized into fewer age groups by averaging the ratios for adjacent age groups, as has been done for the lower panel of this table. The very sharp and very consistent differentials in fertility associated with differences in educational attainment are clearly revealed in this table. The very high ratios for the youngest ages are probably due to selection: girls who became pregnant were forced to drop out of school; young girls with large amounts of education have remained in school; hence, they have only recently married. On the other hand, it should be noted that for the intermediate age group (25-39 years of age), the differentials have decreased in comparison with the older generation. Thus, aside from the special selective effect at the youngest ages, fertility differentials by education seem to be decreasing. This could be verified precisely by computing a table such as 18-24 for a preceding census.

Because CEB data are not highly sensitive to recent fertility changes but are able to measure differential fertility over sustained periods of time, the greatest utility of CEB data is in the study of differential fertility, in addition to their use in cohort analysis.

At this point it is important to note the sources of data for the study of fertility differentials. Conventional vital statistics offer few opportunities. In the U. S. they show only color differentials and area differentials.

The major source of extensive data for current differentials in fertility lies in the exploitation of own-children data, as described in Section 18-13. Using the procedures described there, it is possible to develop current ASFR, GFR, and TFR for education, occupation, income ethnic, migration, housing, or any other set of categories for which the census collects data. These traits may refer to the mother, the father, or both. These measures may be computed for all women, ever married women, or currently married women. GFR may be simultaneously standardized for age and educational attainment (or other variables) when studying a particular differential. Thus, the "own children" technique offers vast potential for the study of fertility differentials. (For an exhaustive exploration of this, the reader may wish to consult the monograph Differential Current Fertility in the United States, by Lee-Jay Cho, Wilson Grabill and Donald J. Bogue (1970). Tabulations of children ever born are, by definition, the sole source of longitudinal or real cohort fertility differentials.

The measurement of fertility differentials, whether cross-sectional or longitudinal, does not explain why they exist. Explanation requires the introduction of additional variables thought to have explanatory value. If the explanatory variable does indeed account for the differential, then the differential will disappear when this explanatory variable is held constant. A widely used explanation of differential fertility is educational attainment. So universal and consistent are educational differentials in fertility and so interrelated with all other variables explanatory of fertility, it is desirable to control the effect of education before assessing the explanatory effects of religion, race, occupation, migratory status, etc. Thus, the ideal measure of differential fertility is one where age, nuptiality, and educational attainment are simultaneously controlled, either through multiple cross-classification or multiple standardization.

Ultimately, the explanation of intergroup differences in age-specific nuptial fertility is the differential efficiency with which groups adopt a practice contraception. This is true because age-specific differences in fecundability are thought to be negligible among most populations, and sexual activity in all cultures is sufficient to assure sustained exposure to pregnancy if contraception is not practiced. Therefore, a goal of the study of differential fertility is that of interpreting the relative acceptance and use of contraception by subgroups of the married population.

18-15. Proportional Fertility or Birth Ratios. Births may be classified into categories that refer to particular traits. Sex of the child or the month in which the birth occurred are obvious examples. Demographers and others often wish to learn how the births are distributed according to the categories of these traits. What proportion of all births are female babies? What proportions of births are born in each month? The need here is not for fertility rates as discussed in previous sections, but only for birth proportions which merely report what proportion of all births fall into a given category of some trait.

$$\text{Birth proportion} = \frac{B_i}{B} \cdot k \quad (37)$$

where  $B_i$  is the number of births in some specific category during a stated time interval,  
 $B$  is the total number of births during the same interval of time, and  
 $k$  is a constant, such as 100 or 1000.

Alternatively, birth ratios may be calculated reporting the ratio of one category of births to another.

$$\text{Birth ratio} = \frac{B_i}{B_j} \cdot k \quad (38)$$

where  $B_i$  and  $B_j$  are the number of births in categories  $i$  and  $j$ .

From the viewpoint of the childbearing female population, equations (37) and (38) are not true rates and are subject to the limitations of all proportional ratios. If the infants themselves are regarded as "exposed" to the likelihood of falling into one of the classes of births, then each ratio may be interpreted as a rate. Birth proportions that are of interest to demographers include:

- a. Sex ratio or sex proportions at birth
- b. Birth-order proportions
- c. Illegitimacy proportion--illegitimate births as a proportion of all births
- d. Birth-weight distribution--proportion of births in each birth-weight class
- e. Duration of pregnancy proportions--proportions of all live births in duration of gestation classes
- f. Hospital birth proportion--proportion of births that occur in hospitals
- g. Physician attendance proportion--proportion of births attended by a physician
- h. Multiple birth proportion--proportion of all pregnancies that are multiple births; proportion of these that are twins, triplets, etc.
- i. Birth composition--proportional distribution of births according to some trait of the parents: age, race, ethnicity, occupation, religion, educational attainment.

The temptation to make inferences about birth rates or changes in birth rates from statistics of birth proportions or changes in birth proportions should be avoided. An example of this fallacy is discussed in connection with birth order statistics.

18-16. Birth Order Proportions and Rates. A great deal of interest is given to birth order in fertility analysis because of its indirect implications for population growth. In bearing a first and second child, parents are merely replacing themselves. Bearing a third, fourth, or higher-order child is a contribution to population growth. The greater the proportion of births that are of high order, the higher the birth rate. The statistical measures which are used to study birth order phenomena range from simple to complex. The denominators of some of these measures require that women be classified by parity, or the number of children they have previously borne. For example, a woman who has borne no children is of zero parity; a

first parity woman has borne one child, etc. Following are the five measures that are most frequently encountered.

1. Birth order proportions--the proportion of all births which are first, second, etc., order:  ${}^k_B/B$ , where  ${}^k_B$  is births of order  $k$ .
2. Order-specific fertility rates (OSFR)--the number of births of order  $k$  per 1000 women of childbearing age:

$$\text{OSFR} = \frac{{}^k_B}{{}_30F_{15}} \cdot 1000 \quad (39)$$

3. Order-age-specific fertility rates (OASFR)--the number of births of order  $k$  to women of age  $x$  per 1000 women of age  $x$ :

$$\text{OASFR} = \frac{{}^k_{B_x}}{F_x} \cdot 1000 \quad (39a)$$

4. Order-parity-specific fertility rates (OPSFR)--the number of births of order  $k$  per 1000 women of childbearing age of parity  $k-1$ :

$$\text{OPSFR} = \frac{{}^k_B}{{}_{k-1}F_{15}} \cdot 1000 \quad (40)$$

5. Age-parity-specific fertility rates (APSFR)--the number of births to women of order  $k$  and age  $x$  per 1000 women of parity  $k-1$  and age  $x$ :

$$\text{APSFR} = \frac{{}^k_{B_x}}{{}_{k-1}F_x} \cdot 1000 \quad (41)$$

Table 18-25 illustrates the computation of these rates.

The above rates may be computed either on a period or a real cohort basis. They may be computed with respect to all women or with respect to married women (nuptial-order/parity-specific rates). OPSFR and APSFR may be calculated either as central rates or as birth probabilities.

Order-parity-specific and age-parity-specific fertility rates are highly useful in the interpretation of fertility change. Their use for this purpose will be discussed in a later chapter. These two rates are not widely used, however, because of the difficulties of getting precise estimates of women of each parity or age-parity.

Table 18-25. ILLUSTRATION OF PROCEDURE FOR CALCULATING ORDER-SPECIFIC, PARITY-SPECIFIC, AND AGE-PARITY SPECIFIC FERTILITY RATES; UNITED STATES: 1960

Age of woman	Total	Number of births by birth order							
		First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eight and over
Total.....	4,257,850	1,090,152	1,022,356	797,402	511,308	292,610	168,210	99,882	154,532
15-19 years.....	593,746	403,076	140,306	32,728	5,830	834	120	34	38
20-24 years.....	1,426,912	471,036	473,104	269,974	114,262	41,796	14,004	4,274	1,694
25-29 years.....	1,092,816	141,442	260,044	272,678	182,198	99,802	53,724	27,324	21,114
30-34 years.....	687,722	50,410	102,510	148,162	130,794	87,950	55,630	35,976	52,688
35-39 years.....	339,908	19,712	38,638	61,392	64,056	49,600	34,832	24,600	54,596
40-44 years.....	96,746	4,496	7,454	12,468	14,168	12,628	9,900	7,624	24,402
Age of woman	Total	Number of women by parity							
Total.....	36,093,401	12,199,570	5,847,131	7,435,241	5,016,983	2,707,005	1,335,456	685,775	866,242
15-19 years.....	6,588,597	6,061,509	415,082	98,829	13,177	....	....	....	....
20-24 years.....	5,519,937	2,550,211	1,407,584	982,549	397,435	126,959	38,640	11,040	5,520
25-29 years.....	5,537,104	1,046,513	1,063,124	1,550,389	1,013,290	492,802	215,947	94,131	80,908
30-34 years.....	6,111,422	855,599	886,156	1,619,527	1,289,510	733,371	354,462	183,343	189,454
35-39 years.....	5,418,536	770,224	1,014,129	1,688,075	1,290,126	757,387	391,531	211,812	295,253
40-44 years.....	5,917,805	935,013	1,059,287	1,485,369	1,029,698	603,616	319,561	183,452	301,809
Age of woman	Total	Age-parity specific fertility rates							
Total.....	89.1	67.4	356.8	364.9	396.7	....	....	....	....
15-19 years.....	258.1	186.8	341.4	281.0	299.8	349.7	408.0	444.4	....
20-24 years.....	197.4	140.2	252.3	181.0	183.7	207.6	258.3	303.2	352.1
25-29 years.....	118.7	61.2	120.1	95.0	104.6	123.7	161.5	201.1	282.4
30-34 years.....	56.2	27.1	39.7	38.1	51.7	68.2	93.9	120.3	195.1
35-39 years.....	15.5	4.8	7.3	8.3	13.6	20.6	30.9	39.4	76.5
40-44 years.....	118.0	90.3	178.0	109.5	104.1	110.6	129.4	147.1	182.1
Parity specific....	118.0	31.1	29.2	22.8	14.6	8.3	7.6	7.6	4.3
Order specific.....									

Age-parity specific rates = births of each order and age divided by women of same age and parity equal to order minus 1.

Order-parity specific rates = total births of each order divided by women of parity equal to order minus 1.

Order specific rates = total births of each order divided by all women 15-44 years of age.

Order-age-specific fertility rates = births of each order and age divided by all women of age x, irrespective of parity

(calculations not shown)

Source: Births from Table 2-13 of Vital Statistics of the United States: 1960. Number of women is census total distributed by percentage distribution of women by parity, Table 1-M, Vital Statistics of the United States: 1960. Rates are as reported in Table 1-N of this report. Data for calculation of rates have been adjusted for underregistration and underenumeration

18-17. Estimation of Basic Fertility Measures by Regression Analysis. Using the knowledge that all indices of fertility are highly intercorrelated (See Section 18-4.), it is possible to employ multiple regression analysis to estimate all of the basic fertility measures, using census data and a few other variables that are easily available or estimated. These equations are easy to understand, easy to use, and yield a surprisingly high degree of precision. The argument is not complicated: Using nations for which valid data on fertility are available, one isolates a group of independent variables (explanatory variables) which may be calculated from the census or other readily available sources and which, when used in appropriate multiple regression equations, can be used to estimate reliably the principal fertility measures--GFR, CBR, TFR and ASFR. The constants which are involved in these multiple regression equations (coefficients of partial regression and intercept) are calculated by using the data for the nations where information both for fertility and for the explanatory variables are available. In order to estimate the fertility rates for a country with poor vital statistics, the values of the explanatory variables are substituted into these equations. The procedure is "theory based" because the only explanatory variables that are used are known to have a causal influence upon fertility levels. The procedure is "empirical" because the multiple regression equation average out place-to-place variations and make partial adjustments for numerous other explanatory factors that cannot be explicitly quantified.

This multiple regression procedure for the estimation of fertility rates was first introduced by Bogue and Palmore in 1964 (Demography, Vol. 1). Two systems of multiple regression equations are presented below. System O (for original) is based upon the original computations of Bogue and Palmore. The data for this system are based upon observations for 50 nations as of 1960. In this set of calculations, major emphasis was placed upon high quality and recency of data. System R (for revised) is based upon a revised system developed by Bogue especially for this monograph. The data for this system are for the same 50 nations, but they include observations for earlier years (as early as 1900) where the data were available and appear to have been valid. The two systems differ in four respects:

- (a) In System O, the data are heavily weighted with observations for populations which have long had low birth rates, while System R contains relatively more observations for nations in intermediate stages of the demographic transition.
- (b) The quality of the data is somewhat higher for System O than for System R because both vital registration and census enumerations were sometimes less exact earlier

in the century than in 1960.

- (c) In System O, GFR refers to women aged 15-49 years, following the United Nations definition of the childbearing ages. In System R, GFR refers to women aged 15-44, as recommended earlier in this chapter.
- (d) The explanatory (independent) variables are not all defined identically in the two sets, and the sequence of making estimates is not identical. Nevertheless, the two systems purport to estimate the same set of fertility measures.

It is recommended that System R be used in preference to System O. If only one set of estimates is to be prepared, however, the amount of calculation for any one nation is so small that a better procedure would be to estimate fertility twice, using each system independently. The results for TFR, CBR, and ASFR should be nearly identical, although they will differ for GFR because of the different age range used in the denominator. If substantially different results are obtained by the two systems, a variable-by-variable comparison should be made to identify the source of the discrepancy. This will almost automatically suggest a preferred estimate because a deficiency in the data for one or more of the explanatory variables usually will be found to be the cause.

The exposition which follows is made primarily in terms of System R. The equations for System O are presented last, with notes that indicate how they are to be used.

18-17a. Estimation of Basic Fertility Measures by Regression Analysis: System R. The principal explanatory (independent) variables of this approach are ratios of children to women, which may easily be calculated from census or survey data. These ratios are known to be highly sensitive indicators of fertility levels. In the system proposed here these ratios are expressed in three forms, each one as a variable.

<u>Symbol</u>	<u>Variable name</u>
$X_{1a}$	Ratio of children 0-4 years of age to women 15-44 years of age
$X_{1b}$	Ratio of children 5-9 years of age to women 20-49 years of age
$X_{1c}$	Ratio of children 0-14 years of age to women 15-49 years of age

Using 87 observations for 50 nations for which valid data from both vital statistics and censuses are available, the close correlation between these three indicators of fertility and the basic fertility may be noted from the data reported below. These observations refer to years in which censuses were taken, from 1900 to 1961. Thus, the

data refer to all stages of the demographic transition; hence, they are appropriate in a wide variety of contexts.

Measure	Correlation with child-woman ratio:		
	$X_{1a}$	$X_{1b}$	$X_{1c}$
Crude birth rate.....	.894	.848	.863
General fertility rate.....	.945	.888	.909
Total fertility rate.....	.942	.897	.915
Age-specific fertility rate:			
15-19 years.....	.537	.490	.479
20-24 years.....	.742	.649	.678
25-29 years.....	.893	.822	.863
30-34 years.....	.806	.786	.800
35-39 years.....	.737	.755	.751
40-44 years.....	.632	.664	.652

With such high correlations it is evident that a reasonably good estimate of each of the above basic measures (except ASFR for the age group 15-19 years) could be made from simple linear equations linking one of the child-woman ratios to fertility. However, it is possible to do much better. By utilizing multiple linear regression equations to introduce a few additional factors which also affect the fertility level, estimates of the basic fertility measures can be made even more precise.

Additional factors which affect fertility are:

Symbol	Variable name
$X_2$	<u>Age at marriage and proportion never marrying</u> --The later the age at marriage and the greater the proportion never marrying, the lower the birth rate tends to be.
$X_3$	<u>Child mortality</u> --The higher the rate of child and infant mortality, the greater the tendency for child-woman ratios to underestimate the true level of fertility.
$X_4$	<u>Female mortality</u> --The higher the rate of maternal mortality, the greater the tendency for child-woman ratios to overestimate the true level of fertility.
$X_5$	<u>Age composition of the women of childbearing age</u> --The greater the concentration of women of childbearing age in the age group 15-29, the higher the level of fertility tend to be.

By the multiple regression approach, these factors may be introduced simultaneously with the ratio of children to women into the regression analysis as additional variables:

Symbol	Variable name
$X_{2a}$	Proportion of women married at age 15-19 years
$X_{2b}$	Proportion of women married at age 20-24 years
$X_{2c}$	Proportion of women married at age 25-29 years
$X_{2d}$	Proportion of women married at age 30-34 years
$X_{2e}$	Proportion of women married at age 35-39 years
$X_3$	Infant mortality rate-- $q_0$ at the time of the census
$X_4$	Survival of women of childbearing age--proportion of women aged 15 who survive to age 40, computed from an appropriate life table
$X_{5a}$	Proportion of all women between ages 15 and 44 who are aged 15-19 years
$X_{5b}$	Proportion of all women between ages 15 and 44 who are aged 20-24 years
$X_{5c}$	Proportion of all women between ages 15 and 44 who are aged 25-29 years
$X_{5d}$	Proportion of all women between ages 15 and 44 who are aged 30-34 years
$X_{5e}$	Proportion of all women between ages 15 and 44 who are aged 35-39 years

All of the above items of data can be obtained from a national census or sample survey or can be estimated from model life tables. Data that are not available may be "borrowed" from other populations thought to be generally similar. It is not necessary to make use of all of the above variables in the multiple regressions. Since many of them are themselves highly intercorrelated, some of them serve as indices of others. By a judicious combination of demographic theory and stepwise regression analysis, it is possible to arrive at equations for each of the basic fertility measures that make use of a minimum number of variables in predicting the respective fertility rates.

A variety of alternative sequences could be employed to develop predictive equations by the above argument. We believe that the one outlined below yields with maximum theoretical justification estimates that are precise and internally consistent. Its primary feature is that estimates are made first for those fertility measures which can be estimated with the greatest precision. At a later stage these estimates are used in connection with the explanatory variables to estimate the remaining fertility measures. The procedure requires the following steps:

- Step 1. Make a preliminary estimate of the general fertility rate.
- Step 2. Make a preliminary estimate of the total fertility rate.
- Step 3. Make preliminary estimates of the age-specific ferti-

ity rates in which the data for each age group are free to vary independently of each other.

- Step 4. Make second estimates of ASFR in which values of some ASFR are used to estimate other ASFR.
- Step 5. Make final estimates of ASFR by averaging the ASFR of Step 3 with the ASFR computed in Step 4.
- Step 6. Make final estimates of GFR, TFR, and CBR on the basis of the schedule of ASFR established in Step 5.

The above sequence of six steps may be carried out in four alternative ways, depending upon which of the child-woman ratios are employed as the principal explanatory variables.

- Set A            uses variable  $X_{1a}$ ;
- Set B            uses variables  $X_{1a}$  and  $X_{1b}$ ;
- Set C            uses variable  $X_{1c}$ ; and
- Set D            uses variables  $X_{1a}$ ,  $X_{1b}$ , and  $X_{1c}$ .

If data on the age of young children are comparatively precise, Set A, B or D is preferred. Set C is presented for use where age data on young children is poor and only broad ratios can be trusted. Research has shown that the ratio of children aged 5-9 to women aged 20-49 used alone is almost always less predictive of fertility than either the ratio of children 0-4 to women 15-44 or the ratio of children 0-14 to women 15-49. Hence, it is not offered alone as one of the options.

The explanatory variables in the equation for each set are submitted in two subsets. Subset 1 (maximum precision) incorporates all relevant variables. Subset 2 (maximum efficiency) incorporates only those variables for which the standard error of the standardized regression coefficient (beta) is equal to or smaller than the standardized regression coefficient. The equations that comprise each of the four sets are presented in Table 18-26 through 18-31. Following is a description of the argument by which Set A is developed. The same principles were used in developing each of the sets; hence, it would be superfluous to make a detailed discussion of each. Instead, a few supplementary comments are provided that identify the principles governing choice of variables for these other sets.

Step 1. Establish a Preliminary Estimate of the General Fertility Rate. The general fertility rate can be estimated by multiple regression equations with a smaller relative error than any of the other basic fertility measures. Hence, it provides the best foundation for the system. It is affected by all of the variables listed above. Subset 1 reports the regression constants needed to estimate GFR using all of the explanatory variables. However, the constants for

Table 18-26. EQUATIONS FOR ESTIMATING GENERAL FERTILITY RATE FROM EXPLANATORY VARIABLES DERIVED FROM CENSUSES AND OTHER SOURCES, SYSTEM R

Explanatory variables	Symbol in regression equation	Constants for multiple-regression estimation equations							
		Set A		Set B		Set C		Set D	
		Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2
<u>Partial regression coefficients</u>									
<u>A. Child-woman ratios</u>									
Children 0-4 to women 14-44 years.....	X <sub>1a</sub>	224.1	226.8	189.4	207.0	....	....	184.9	207.0
Children 5-9 to women 20-49 years.....	X <sub>1b</sub>	....	....	50.0	36.6	....	....	42.7	36.6
Children 0-14 to women 15-49 years.....	X <sub>1c</sub>	....	....	....	....	118.4	119.3	5.7	....
<u>B. Proportions married</u>									
Women aged 15-19 years.....	X <sub>2a</sub>	37.6	24.6	28.6	....	25.2	....	28.9	....
Women aged 20-24 years.....	X <sub>2b</sub>	- 6.1	....	- 6.3	....	3.9	....	- 6.1	....
Women aged 25-29 years.....	X <sub>2c</sub>	23.4	....	28.6	....	46.3	47.8	28.9	....
Women aged 30-34 years.....	X <sub>2d</sub>	8.3	....	35.9	....	238.5	273.6	40.7	....
Women aged 35-39 years.....	X <sub>2e</sub>	- 49.6	....	- 94.2	....	-392.4	-422.9	-101.4	....
<u>C. Mortality</u>									
Infant mortality rate.....	X <sub>3</sub>	.1	.1	.1	.1	.1	.2	.1	.1
Female survival, age 15 to 40.....	X <sub>4</sub>	14.3	....	36.7	....	72.0	72.6	35.9	....
<u>D. Age composition of women</u>									
15-44									
Proportion aged 15-19 years.	X <sub>5a</sub>	43.5	....	- 57.6	- 83.6	-292.5	-273.8	- 63.2	- 83.6
Proportion aged 20-24 years.	X <sub>5b</sub>	202.5	223.4	237.4	253.9	347.7	337.7	239.8	253.9
Proportion aged 25-29 years.	X <sub>5c</sub>	65.0	....	13.0	....	- 68.5	....	11.4	....
Proportion aged 30-34 years.	X <sub>5d</sub>	-214.0	-194.6	-232.5	-211.0	-289.3	-302.1	-233.5	-211.0
Proportion aged 35-39 years.	X <sub>5e</sub>	222.7	215.9	177.0	216.8	- 7.7	....	173.9	216.8
Intercept with the regression line.....		-59.650	-54.661	-44.373	-48.427	44.335	29.343	-41.441	-48.427
Explained variance (RHO squared).....		.932	.930	.934	.931	.905	.904	.935	.931
Standard error of estimate.....		8.962	8.662	8.826	8.632	10.530	10.321	8.862	8.632

Source: Based on 87 observations for 50 nations with reliable data for selected census years, 1900 to 1960-61.

Notes: (1) Subset 1 makes use of all explanatory variables in the system, while Subset 2 makes use only of explanatory variables where the standard error of the regression coefficient is equal to or smaller than the regression coefficient itself.  
(2) The symbol (....) indicates that the particular variable is not involved in the equation for this set or subset.

Subset 2 demonstrate that an almost equally precise estimate can be constructed by using only one or two representatives of each factor. As Table 18-26 shows, this may be done for each of the 4 sets with a maximum of four or five variables. To write the regression equations it is necessary merely to substitute the partial regression coefficients and the intercept into the equation, as the following example for Set A illustrates:

$$\begin{aligned} \text{GFR}(\text{Set A}) = & 224.1 X_{1a} + 37.6 X_{2a} - 6.1 X_{2b} + 23.4 X_{2c} + 8.3 X_{2d} \\ & - 49.6 X_{2e} + 0.1 X_3 + 14.3 X_4 + 43.5 X_{5a} + 202.5 X_{5b} \\ & + 65.0 X_{5c} - 214.0 X_{5d} + 222.7 X_{5e} - 59.65 \end{aligned}$$

Step 2. Establish a Preliminary Estimate of the Total Fertility Rate. The total fertility rate is somewhat more variable than the general fertility rate; therefore it has a larger relative standard error when calculated directly from child-woman ratios and other variables. Moreover, for the most precise estimate of TFR we need an estimate that has maximum consistency with the GFR. Therefore, we develop equations which include the GFR estimated in Step 1 as one of the explanatory variables, along with additional variables that improve the precision of the estimate. The four sets of equations recommended for this are reported in Table 18-27. The value of TFR estimates by these equations will be highly consistent with the estimated value of GFR, which we obtained previously. In fact, it will be almost 100 percent consistent with this estimate. This is demonstrated by values of  $R^2$  equal to .997 for all equations of Table 18-27. If, however, an error has been made in the estimation of GFR, this error will be reflected in the estimate of TFR. Inasmuch as GFR is the average fertility rate of childbearing over 30 years (from age 15 to age 44) one might expect TFR to be exactly equal to 30 times GFR. The first line of coefficients of Table 18-27 demonstrates that this is very nearly the case. However, by introducing additional factors, an estimate that is even more precise than this formal definition is attainable.

There may be instances where it is desired to estimate TFR directly, without using the estimated value of GFR as one of the explanatory variables. A set of regression equations for accomplishing this is reported in Table 18-28. Usually it would be advisable to estimate TFR both by using estimated values of GFR (Table 18-27) and by using only the explanatory variables (Table 18-28), and then to compare the results. If there is a substantial discrepancy, its cause should be determined, and a decision should be made concerning the best estimate of TFR before continuing the steps.

Step 3. A Preliminary Estimate of Age-Specific Fertility Rates. Age-specific fertility rates are relatively much more variable than

Table 18-27. EQUATIONS FOR ESTIMATING TOTAL FERTILITY RATE FROM DATA FOR ESTIMATED GFR AND EXPLANATORY VARIABLES DERIVED FROM CENSUS AND OTHER SOURCES, SYSTEM R

Explanatory variables	Symbol in regression equation	Constants for multiple-regression equations							
		Set A		Set B		Set C		Set D	
		Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2
<u>Partial regression coefficients</u>									
<u>A. General fertility rate.....</u>									
A. <u>Child-woman ratios</u>	$X_G$	29.3	29.4	29.0	29.1	30.1	30.0	29.1	29.1
Children 0-4 to women									
15-44 years.....	$X_{1a}$	4.3	3.9	3.0	2.6	....	....	6.6	5.6
Children 5-9 to women									
20-49 years.....	$X_{1b}$	....	....	2.6	3.0	....	....	8.6	8.8
Children 0-14 to women									
15-49 years.....	$X_{1c}$	....	....	....	....	1.1	1.5	- 4.7	- 4.1
<u>B. Proportions married</u>									
Women aged 15-19 years.....	$X_{2a}$	- 9.0	- 9.0	- 9.4	- 9.6	- 9.3	- 9.4	- 9.7	-10.1
Women aged 20-24 years.....	$X_{2b}$	- .2	....	- .2	....	....	....	- .4	....
Women aged 25-29 years.....	$X_{2c}$	4.0	3.9	4.3	4.2	4.0	4.0	4.0	3.9
Women aged 30-34 years.....	$X_{2d}$	-16.3	-16.2	-14.8	-14.9	-13.7	-14.2	-18.8	-18.8
Women aged 35-39 years.....	$X_{2e}$	13.6	13.4	11.1	11.1	10.1	10.7	17.1	16.8
<u>C. Mortality</u>									
Infant mortality rate.....	$X_3$	- .0	- .0	- .0	- .0	- .0	- .0	- .0	- .0
Female survival, age 15 to 40.....	$X_4$	-10.0	-10.1	- 8.8	- 8.8	- 9.5	- 9.6	- 8.2	- 8.2
<u>D. Age composition of women 15-44</u>									
Proportion aged 15-19 years.	$X_{5a}$	7.6	6.0	2.4	....	4.8	....	7.1	....
Proportion aged 20-24 years.	$X_{5b}$	-32.6	-35.6	-30.3	-31.4	-33.0	-31.3	-32.4	-32.4
Proportion aged 25-29 years.	$X_{5c}$	-32.0	-34.5	-34.6	-37.1	-36.9	-36.9	-33.3	-39.1
Proportion aged 30-34 years.	$X_{5d}$	-31.7	-31.6	-33.2	-33.5	-30.7	-32.0	-32.3	-33.7
Proportion aged 35-39 years.	$X_{5e}$	5.7	....	3.8	....	1.4	....	6.3	....
<u>Intercept with the regression line.....</u>									
		219.9	244.3	226.6	244.2	234.8	248.2	202.6	236.9
<u>Explained variance (RHO squared).....</u>									
		.997	.997	.997	.997	.997	.997	.997	.997
<u>Standard error of estimate.....</u>									
		5.970	5.913	5.931	5.822	6.063	6.002	5.740	5.674

Source: Based on 87 observations for 50 nations with reliable data for selected census years, 1900 to 1960-61.

- Notes: (1) Subset 1 makes use of all explanatory variables in the system, while Subset 2 makes use only of explanatory variables where the standard error of the regression coefficient is equal to or smaller than the regression coefficient itself.
- (2) The General Fertility Rate is estimated by one of the equations reported in Table 18-20, and used here as one of the explanatory variables.
- (3) The symbol (....) indicates that the particular variable is not involved in the equation for this set or subset.

Table 18-28. EQUATIONS FOR ESTIMATING TOTAL FERTILITY RATE DIRECTLY FROM EXPLANATORY VARIABLES DERIVED FROM CENSUS AND OTHER SOURCES, SYSTEM R

Explanatory variables	Symbol in regression equation	Constants for multiple-regression equations							
		Set A		Set B		Set C		Set D	
		Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2
Partial regression coefficients									
A. Child-woman ratios									
Children 0-4 to women 15-44 years.....	X <sub>1a</sub>	6986.6	6961.9	5794.2	5819.7	....	....	6035.3	5819.7
Children 5-9 to women 20-49 years.....	X <sub>1b</sub>	....	....	1716.7	1510.5	....	....	2107.4	1510.5
Children 0-14 to women 15-49 years.....	X <sub>1c</sub>	....	....	....	....	3678.1	3681.2	- 305.8	....
B. Proportions married									
Women aged 15-19 years.....	X <sub>2a</sub>	203.1	....	- 108.3	....	- 179.0	....	- 128.2	....
Women aged 20-24 years.....	X <sub>2b</sub>	- 198.2	....	- 204.8	....	118.8	....	- 214.5	....
Women aged 25-29 years.....	X <sub>2c</sub>	1085.6	....	1264.3	1082.5	1798.3	1780.4	1245.9	1082.5
Women aged 30-34 years.....	X <sub>2d</sub>	- 1387.8	....	- 439.4	....	5808.6	6181.6	- 699.6	....
Women aged 35-39 years.....	X <sub>2e</sub>	- 89.1	....	- 1620.5	- 2203.4	- 10795.9	- 10924.4	- 1238.9	- 2203.4
C. Mortality									
Infant mortality rate.....	X <sub>3</sub>	2.9	2.8	3.4	3.2	4.0	3.2	3.4	3.2
Female survival, age 15 to 40.....	X <sub>4</sub>	- 585.8	- 1033.1	182.9	....	1215.1	....	225.8	....
D. Age composition of women									
15-44									
Proportion aged 15-19 years.	X <sub>5a</sub>	2037.1	1924.8	- 1432.8	....	- 8327.5	- 8594.2	- 1129.3	....
Proportion aged 20-24 years.	X <sub>5b</sub>	2666.2	....	3863.5	3033.4	7167.5	6755.0	3733.8	3033.4
Proportion aged 25-29 years.	X <sub>5c</sub>	- 1296.8	....	- 3082.1	....	- 5394.0	- 5631.9	- 2996.0	....
Proportion aged 30-34 years.	X <sub>5d</sub>	- 9436.2	- 9422.5	- 10073.9	- 10360.8	- 11777.3	- 11755.9	- 10021.5	- 10360.8
Proportion aged 35-39 years.	X <sub>5e</sub>	7093.5	6150.8	5523.7	6413.8	- 91.7	....	5690.1	6413.8
Intercept with the regression line.....									
		453.3	791.5	977.8	680.8	3682.0	4860.9	820.7	680.8
Explained variance (RHO squared).....									
		.932	.928	.936	.935	.902	.900	.936	.935
Standard error of estimate.....									
		26.894	26.324	26.299	25.459	32.275	31.610	26.462	25.459

either GFR and TFR. Hence, they are more difficult to estimate. Moreover, they are much less sensitive to the child-woman ratios and much more sensitive to data on marital status and age distribution than are GFR and TFR. In addition they are highly responsive to the overall level of fertility, as indicated by GFR and TFR. A first estimate of ASFR, therefore, consists of a set of equations in which both GFR and TFR (as estimated by Steps 1 and 2) are incorporated as explanatory variables, along with some or all of the other variables. Table 18-29 presents the data for writing this set of equations. In this first estimate of ASFR, the rate for each age group is free to vary independently of the values obtained for other age groups. Despite the use of many variables, the estimates for ages 15-19 and 40-44 are subject to larger relative errors than are estimated ASFR for other age groups. Yet even at these extreme ages the value of R is above .90 and the effect of errors in estimating these age groups upon TFR is very small.

The reader will note that in Table 18-29 we have dropped the child-woman ratios as explanatory variables. Instead, we rely entirely upon the estimated values of GFR and TFR to establish the general fertility level.

Step 4. A Second Estimate of Age-Specific Fertility Rates. A second set of estimating equations for ASFR may be established in which the values for some ASFR are used as explanatory variables in estimating the values for other age groups. By this procedure it is possible to improve the estimates for the youngest and oldest age groups, which are subject to the greatest relative error by the multiple regression procedure outlined in Step 3. When this procedure is carried out, the researcher selects one age group that may be estimated most reliably using estimates of GFR, TFR and the explanatory variables. He then uses the estimated ASFR for this age group as an explanatory variable in estimating ASFR for the second-most-easy-to-estimate age group. By successively adding estimates for ASFR as explanatory variables, in order of their size and stability, a refined estimate for each age group is developed.

A sequence of estimation recommended for this procedure is as follows:

Sequence in which ASFR is to be estimated	Indices of fertility level to be used in addition to estimated GFR and TFR
Age 30-34.....	Child-woman ratio of children 0-14 to women 15-49 years
Age 25-29.....	ASFR 30-34
Age 20-24.....	ASFR 30-34, 25-29
Age 35-39.....	ASFR 30-34, 25-29, 20-24
Age 15-19.....	ASFR 30-34, 25-29, 20-24, 35-39
Age 40-44.....	ASFR 30-34, 25-29, 20-24, 35-39, 15-19



Table 18-30 presents the multiple regression constants for a system of equations that follows the above sequence. We see that the standard error of the estimate for the youngest and oldest ages is reduced greatly in comparison with the results reported for Step 3.

Step 5. Final Estimate of Age-Specific Fertility Rates. Although the estimates of ASFR obtained in Step 4 clearly have smaller errors of estimation than the results of the procedure outlined in Step 3, this precision is attained by limiting the freedom of independent variation at the youngest and oldest age groups. When estimating the fertility of a nation whose true fertility is unknown, the equations of Step 4 will produce estimates that are highly consistent; but these estimates may be biased by initial errors of estimating values for ASFR for ages 30-34 and 25-29. It is recommended that a final estimate of ASFR be derived by a simple average of the estimates of ASFR obtained in Steps 3 and 4. This should produce estimates in which consistency is high and yet in which each age group has been at least partially free to vary independently of the values for other age groups.

Step 6. Final Estimates for TFR, GFR, and CBR. By summing the ASFR obtained in Step 5 and multiplying by 5, a final value for TFR is obtained. By applying the values of ASFR to the age distribution of the female population to obtain expected births and then dividing expected births by the total number of women aged 15-44, a final value of GFR is obtained. The crude birth rate may be established by applying GFR or ASFR to the census data or another estimate of the population to obtain an estimate of the total number of births that would be expected in one year. This total is then divided by the total population and the quotient is multiplied by 1000. For maximum precision it is recommended that the schedule of ASFR be used for this purpose.

A Note on the Quality of the Data. It must be kept in mind that the regression equations reported here are all based on data that are thought to be reasonably accurate. If these equations are used to estimate fertility rates for nations where the census data are deficient, it is necessary to adjust the census data for known or suspected errors before substituting into the regression equations. This is especially important for nations of Latin America, where data for marital status may fail to include populations living in consensual unions, and all censuses where children may be seriously underenumerated or where age misstatement may lead to serious errors in the ratios of children to women. Where data are of doubtful quality it would be wise to estimate fertility by all four sets of equations A, B, C and D, and then to compare the results.

Comparison of the Precision of Regression Estimates with Estimates Derived by Other Techniques. In general, the regression estimates derived by the procedures described above will be about as

Table 18-30. EQUATIONS FOR ESTIMATING AGE-SPECIFIC FERTILITY RATES FROM DATA FOR ESTIMATED GFR AND TFR AND SELECTED OTHER ASFR PLUS EXPLANATORY VARIABLES DERIVED FROM CENSUSES AND OTHER SOURCES, SYSTEM R

Explanatory variables	Symbol in regression equation	Constants for multiple regression equations											
		Age 15-19 years		Age 20-24 years		Age 25-29 years		Age 30-34 years		Age 35-39 years		Age 40-44 years	
		Subset 1	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2	Subset 1*	Subset 2	Subset 1	Subset 2	Subset 1	Subset 2
Partial regression coefficients													
General fertility rate.....	X <sub>G</sub>	.731	.644	2.222	1.851	2.886	3.207	-	.624	-	.242	.273	.177
Total fertility rate.....	X <sub>T</sub>	.094	.095	.013	...	.051	.063		.073		.074	.198	.200
A. Age-specific fertility rates													
Women 15-19 years.....	X <sub>a1</sub>	...	...	...	...	...	...		...		...	1.051	1.045
Women 20-24 years.....	X <sub>a2</sub>	-.372	.357		...	...	...		...	-.515	...	1.001	1.032
Women 25-29 years.....	X <sub>a3</sub>	-.820	.822	.518	.540	...	...		...	-.301	...	1.044	1.018
Women 30-34 years.....	X <sub>a4</sub>	-.117	.078	- 1.088	- 1.124	.216	.236		...	.093	...	1.083	1.076
Women 35-39 years.....	X <sub>a5</sub>	- 1.324	- 1.330	...	...	...	...		...	...	...	1.000	1.013
Women 40-44 years.....	X <sub>a6</sub>	...	...	...	...	...	...		...	...	...	...	...
B. Proportions married													
Women aged 15-19 years.....	X <sub>2a</sub>	94.176	82.410	-117.617	- 94.288	-159.665	-173.244	107.293	97.403	- 74.894	- 86.066	9.435	...
Women aged 20-24 years.....	X <sub>2b</sub>	- 30.723	- 22.709	184.308	176.826	- 74.766	- 68.758	-208.778	-217.464	52.686	64.378	- 17.791	...
Women aged 25-29 years.....	X <sub>2c</sub>	13.971	...	- 29.208	- 33.759	78.402	79.130	87.920	96.420	- 15.052	- 15.685	- 8.524	...
Women aged 30-34 years.....	X <sub>2d</sub>	- 25.088	...	-126.747	-105.299	-185.825	-195.313	-131.514	...	- 59.578	- 37.443	- 21.583	...
Women aged 35-39 years.....	X <sub>2e</sub>	13.282	...	19.278	...	368.755	375.315	173.924	...	39.626	...	33.119	...
C. Mortality													
Infant mortality rate.....	X <sub>3</sub>	.045	.038	.006	...	.065	...	.036	...	.024	.020	.020	.025
Female survival, age 15-40..	X <sub>4</sub>	6.256	...	- 19.781	- 24.209	46.983	...	- 58.660	- 75.719	7.070	...	- 10.603	...
D. Age composition of women 15-44													
Proportion aged 15-19 years.	X <sub>5a</sub>	- 15.528	...	81.710	...	- 6.159	...	7.880	- 87.401	- 23.192	...	- 3.241	...
Proportion aged 20-24 years.	X <sub>5b</sub>	- 38.322	...	-533.319	-503.298	-256.010	-309.295	- 49.695	...	52.088	...	- 24.755	...
Proportion aged 25-29 years.	X <sub>5c</sub>	-247.201	-219.564	31.501	...	-604.672	-643.430	220.167	...	- 69.934	- 48.419	- 77.709	- 54.264
Proportion aged 30-34 years.	X <sub>5d</sub>	- 22.014	...	-445.975	-427.668	35.590	...	2.274	...	69.893	...	9.551	...
Proportion aged 35-39 years.	X <sub>5e</sub>	- 25.207	...	78.718	...	35.010	...	98.155	...	- 90.657	- 78.438	20.979	23.208
Intercept with the regression line.....		54.220	37.285	205.953	237.501	- 75.759	- 1.964	- 37.116	75.741	10.220	40.810	11.870	5.494
Explained variance (RHO squared).....		.961	.961	.985	.984	.959	.958	.974	.970	.993	.992	.992	.991
Standard error of estimate.....		4.665	4.463	7.087	6.910	10.738	10.498	8.454	8.422	3.946	3.835	2.758	2.648

Source: Based on 87 observations for 50 nations with reliable data for selected census years, 1900-1961.

Notes: (1) Subset 1 makes use of all explanatory variables in the system, while Subset 2 makes use only of explanatory variables where the standard error of the regression coefficient is equal to or smaller than the regression coefficient itself.

(2) The General Fertility Rate is estimated by one of the equations reported in Table 18-20, and the Total Fertility Rate is estimated by one of the equations reported in Table 18-21. The estimated value of GFR and TFR may have been adjusted by the researcher before substitution into the equations of this table.

\*Ratio of children 0-4 to women 15-44 is 42.273; ratio of children 5-9 to women 20-49 is -42.841; and ratio of children 0-14 to women 15-49 is 2.555.

Table 18-31. ILLUSTRATION OF THE USE OF REGRESSION EQUATIONS, SYSTEM R, TO ESTIMATE FERTILITY MEASURES FOR CEYLON, 1963

Explanatory variables	Values of independent variable for Ceylon	Regression constants								
		GFR	TFR	Age-specific fertility rates ASFR						
				15-19	20-24	25-29	30-34	35-39	40-44	
<u>Partial regression coefficients</u>										
<u>A. Child-woman ratios</u>										
Children 0-4 to women										
15-44 years.....	.72	....	5819.7	....	....	....	....	....	....	....
Children 5-9 to women										
20-49 years.....	.72	....	1510.5	....	....	....	....	....	....	....
Children 0-14 to women										
15-49 years.....	1.74	118.4	....	....	....	....	....	....	....	....
<u>B. Proportions married</u>										
Women aged 15-19 years.....	.243	25.2	....	136.41	-295.75	-141.83	97.82	128.25	....	....
Women aged 20-24 years.....	.676	3.9	....	108.12	348.52	-121.65	-207.89	-110.16	....	....
Women aged 25-29 years.....	.872	46.3	1082.5	-54.53	-79.25	103.25	92.85	4.78	....	-55.04
Women aged 30-34 years.....	.925	238.5	....	126.74	-124.78	-199.20	-100.57	57.79	....	....
Women aged 35-39 years.....	.946	-392.4	-2203.4	-208.81	87.97	383.24	125.20	-113.17	....	....
<u>C. Mortality</u>										
Infant mortality rate.....	71.0	.1	3.2	-.03	-.00	.04	.04	.01	-.07	....
Female survival, age 15 to 40.....	.898	72.0	....	3.43	42.88	....	-39.28	-30.27	....	....
<u>D. Age composition of women 15-44</u>										
Proportion aged 15-19 years.	.201	-292.5	....	90.99	153.66	....	-76.99	-102.65	....	....
Proportion aged 20-24 years.	.220	347.7	3033.4	-188.66	-639.16	-356.43	-27.34	457.56	648.20	....
Proportion aged 25-29 years.	.200	-68.5	....	-96.67	-449.99	-594.13	172.60	348.74	517.46	....
Proportion aged 30-34 years.	.139	-289.3	-10360.8	-253.07	-414.01	....	-14.18	272.16	382.27	....
Proportion aged 35-39 years.	.144	-7.7	6413.8	80.55	51.89	....	46.36	-126.79	....	....
<u>Intercept with the regression line.....</u>										
Estimated rate.....		44.34	680.8	135.50	183.88	-10.58	-17.57	-62.16	-238.18	....
Officially reported rate.....		183	5196	86	243	297	219	160	66	....
Ratio of estimated to official.....		.97	.98	1.26	.95	1.01	.90	1.07	1.50	....

Note: Adjusted TFR by summing up ASFR = 5355  
Adjusted GFR by dividing new TFR by 29 will be 185.

precise as the reverse survival of age distributions. The values of ASFR derived by this procedure will be more precise than can be obtained by any method other than the "own children" procedure of Section 18-13. On one hand, the regression estimates will handle the factor of mortality in a much more approximate way than does the reverse survival procedure. On the other hand they will take into account (directly and indirectly) other factors that affect fertility estimates: marital status, age distribution, census errors, and cohort changes during declining fertility.

Because the regression equations are based upon censuses where underenumeration is thought to be relatively small, the researcher may wish to inflate his estimates of fertility by 1 to 5 (or even more) percent to make allowance for suspected net underenumeration of children.

Where the data for own children are collected and coded reliably, regression estimates will, in general, be somewhat less precise than own children estimates. Where a substantial proportion of children live away from home and where there is considerable laxity in identifying the parents of children, it is possible that the regression estimates will be the more precise.

One outstanding virtue of the regression estimates is that very little computation is required to make them. The estimates for a country can be worked out on a desk calculator in less than one hour's time. Because of this fact, it is advisable to include this procedure along with others as an additional set of evidence, even if the final results are not accepted as the "best estimate" of actual fertility.

18-17b. Estimation of Basic Fertility Measures by Regression Analysis: System O. As mentioned above, the data upon which these equations are based are rather precise observations as of 1960, where as system R contains somewhat less precise observations for earlier censuses. This system is more heavily weighted with data for populations well advanced in the demographic transition than is true for System R.

The original system of regression equations presented by Bogue and Palmore in 1964 (here called System O) contained fewer variables than System R, but covered essentially the same topics. GFR in System O is expressed in terms of women 15-49 years of age. This system will be found helpful when working with this longer range of reproductive ages.

The explanatory variables employed by this system are:

- $X_{1a}$  Number of children aged 0-4 per 1000 women aged 15-49
- $X_{1b}$  Number of children aged 5-9 per 1000 women aged 15-49
- $X_{1c}$  Number of children 0-14 years of age per 1000 total population

$X_{2a}$	Median age at first marriage
$X_{2b}$	Percent of women 15-19 years of age who have ever married
$X_{2c}$	Percent of women 20-24 years of age who have ever married
$X_{2d}$	Percent of women 25-29 years of age who have ever married
$X_{2e}$	Percent of women 45-49 years of age who have ever married
$X_3$	Infant mortality rate (deaths under 1 year of age per 1000 live births)
$X_4$	Index of fertile age composition. This is a ratio formed by dividing (a) the number of births a particular population's female age distribution (converted to a "standard million") would yield if it had the age-specific fertility rates for the world by (b) the number of births the world's female age distribution (converted to a "standard million") yields when the world's age-specific fertility rates are applied. For the above calculations the estimated age-specific fertility rates for the world were: 78.5, 246.0, 283.3, 212.3, 132.1, 56.9, and 10.9. The estimated "standard million" for the world's females 15-49 years of age was: 195,827, 174,739, 153,801, 139,763, 125,761, 111,878, and 98,231.

Thus, the age composition of women was represented by a single summary measure instead of a separate variable for each 5-year age group as in System R.

Several different strategies for estimation of GFR, TFR, ASFR, and CBR from the above variables are presented in the original article. The following 10-step sequence represents one of the most precise of the alternatives and is presented as representative of the system.

Step 1. Estimate TFR, using the three following regression equations:

$$\text{TFR} = 7.6193 X_{1a} + 11.319 X_3 - 44.963 X_{2a} - 4928.4 X_4 \\ - 7.2189 X_{2e} + 5804.3$$

$$\text{TFR} = 9.2547 X_{1b} + 10.526 X_3 - 98.797 X_{2a} - 3785.1 X_4 \\ - 26.535 X_{2e} + 7125.8$$

$$\text{TFR} = 170.54 X_{1c} + 7.1342 X_3 - 81.818 X_{2a} - 6005.7 X_4 \\ - 16.632 X_{2e} + 6958.3$$

Step 2. Using the estimate of TFR of Step 1, estimate ASFR for age group 25-29 years, using the following equation:

$$\text{ASFR}_{25-29} = 0.0495 \text{TFR} - 0.1925 X_3 + 0.9758 X_{2d} + 7.6215 X_{2a} \\ - 13.380 X_4 - 188.23$$

Step 3. Using the information of Steps 1 and 2, estimate ASFR for the age group 20-24 years, using the following equation:

$$\text{ASFR}_{20-25} = 0.0524 \text{ TFR} - 0.3706 \text{ ASFR}_{25-29} - 0.6431 X_3 - 15.242 X_{2a} \\ + 0.1720 X_{2c} + 431.66$$

Step 4. Use the information of Steps 1 through 3 to get the age-specific rate for the age group 15-19 years, using the following equation:

$$\text{ASFR}_{15-19} = 0.0365 \text{ TFR} + 0.4470 \text{ ASFR}_{20-24} - 0.7426 \text{ ASFR}_{25-29} \\ + 2.4585 X_{2a} + 0.6949 X_{2b} - 76.631$$

Step 5. Use the information of Steps 1 through 4 to get the age-specific rate for the age group 35-39 years, using the following equation:

$$\text{ASFR}_{35-39} = 0.0815 \text{ TFR} - 0.4821 \text{ ASFR}_{15-19} - 0.3235 \text{ ASFR}_{20-24} \\ - 0.5225 \text{ ASFR}_{25-29} + 0.1529 X_{2a} - 0.1674 X_{2e} - 0.5904$$

Step 6. Use the information for adjacent age groups to estimate the rate for ages 30-34, using the following equation:

$$\text{ASFR}_{30-34} = 0.1350 \text{ TFR} - 0.5457 \text{ ASFR}_{15-19} - 0.7687 \text{ ASFR}_{20-24} \\ - 0.5223 \text{ ASFR}_{25-29} - 0.7804 \text{ ASFR}_{35-39} - 0.1776 X_3 \\ - 1.0513 X_{2a} - 0.1227 X_{2e} + 51.734$$

Step 7. Estimate ASFR for ages 40-44, using the following equation:

$$\text{ASFR}_{40-44} = 0.1460 \text{ TFR} - 0.7729 \text{ ASFR}_{15-19} - 0.6583 \text{ ASFR}_{20-24} \\ - 0.8306 \text{ ASFR}_{25-29} - 0.6515 \text{ ASFR}_{30-34} - 0.6873 \text{ ASFR}_{35-39} \\ + 1.1703 X_{2a} + 0.0242 X_{2e} - 28.049$$

Step 8. Estimate ASFR for ages 45-49, using the following equation.

$$\text{ASFR}_{45-49} = - 0.0038 \text{ TFR} + 0.0612 \text{ ASFR}_{25-29} + 0.0330 \text{ ASFR}_{30-34} \\ + 0.1047 \text{ ASFR}_{35-39} - 0.1118 \text{ ASFR}_{40-44} + 0.1065 X_3 \\ - 0.2206 X_{2a} + 0.0457 X_{2e} - 8.2324$$

Step 9. Adjust the age-specific fertility rates obtained in Steps 2 through 8 so that they yield the same total fertility rate as that predicted in Step 1. This is done by taking the ratio:

$$\text{Adjustment factor} = \frac{\text{Originally predicted TFR (Step 1)}}{\text{TFR found by summing ASFR (Steps 2-8)}}$$

and using it as a multiplying adjustment factor on each of the estimated ASFR.

Step 10. Estimate GFR and CBR by applying the adjusted ASFR of Step 9 to the age composition of the population to get expected births. Then take the ratio of expected births to women of child-bearing age. Calculate CBR by taking the ratio of expected births to total population.

As we indicated for System R, the above equations assume that the data for the independent variables substituted into the equations are of high quality. Where the census data (especially ratios of children to women) are deficient, the resulting estimates will be biased accordingly. Hence, it is essential that corrective adjustments be made to these variables before making the substitution. At Step 1, it is especially advisable to experiment with all three equations, and with various plausible adjustments, before proceeding to carry out the remaining steps.

18-18. Deficiencies in Vital Statistics and Census Data for Measuring Fertility and Techniques for Computing Correction Factors. Throughout this chapter it has been noted that vital statistics and census data for measuring fertility may have deficiencies. The various equations that have been presented have contained coefficients for correcting the data for these inadequacies, but the procedure for estimating these correction factors was deferred. In this section we study three of these deficiencies and techniques of adjusting dates to correct for them:

- a. underregistration of children;
- b. underenumeration of children at the census or in sample surveys;
- c. underreporting of children ever born at the census or in sample surveys.

This discussion and these illustrations are in terms of the United States population. The techniques, however, are general and can be extended to other populations. The materials deal primarily with the United States, principally because much intensive research into the extent and sources of error for these data has been made for this nation. The tests that have been made and the results obtained are described in some detail for two reasons. First, they provide a necessary perspective from which to apply correction factors and, second, they may serve as a first introduction to the subject for demographers from other nations who wish to conduct similar tests in their own countries.

18-18a. Underregistration of Births. The basic method of testing completeness of birth registration is to secure in some way a list of children born during a certain period in a given area, and then to search the files of the registrar to determine whether or not a certificate can

be located for each child. The completeness of registration is indicated by the proportion of children for whom a certificate can be found, after making certain that all missing certificates should be found in that particular file.

It has been known for many years that a significant proportion of births in the United States was not being registered, even after a state had been admitted to the birth registration area. Since a completeness of only 90 percent was required before admission to the registration area, the possibilities for undercount were quite large. Even before the beginning of the 20th century, efforts were being made to estimate the magnitude of the error. One procedure consisted of distributing postal cards to a large universe of families, requesting that the card be used to report every child in the family born during the preceding year. The cards returned were checked against birth records. Other attempts included efforts to check birth files against births reported in newspapers and in baptismal records of churches. Tests, such as that of Hedrich and his collaborators (listed in the basic readings), showed that the results of such efforts were biased in such way as to indicate a higher percentage of registration completeness than was actually the case. From field studies it was found that birth registration completeness was higher than average among the prosperous, intelligent, urban, and white families, and for births that occurred in hospitals. These were also the families who were most inclined to return the postcards properly filled out. Registration was found to be especially poor for unattended births and rural Negro births. The investigators found that a high proportion of this group was not accustomed to receiving mail of any kind, and that mail carriers tended not to deliver the postal inquiries at all to such families.

The problem of underregistration of births was also encountered early in efforts to construct life tables, for births were needed in order to arrive at a reliable population to use as a base for mortality rates in the first years of life. Elbertie Foundray made pioneering studies in the construction of the 1919-20 life tables. By first making a rough estimate of the magnitude of the census undercount of infants, she compared registered births with infants counted by the census, allowing for infant deaths between the time of birth and the census count. In this way she obtained estimates of underregistration for the nation as a whole. Improvising upon her estimating procedure, P. K. Whelpton demonstrated in 1934 that birth registration was much less complete in some states than in others, and that it was more complete for white than for Negro births. His studies also indicated that most complete registration was found in those states where registration procedures had been in effect for several decades. It was reasoned that as the registration system continued, physicians, hospitals, midwives

and parents gradually become aware of their responsibility, familiar with the procedure, and sympathetic to the need for registration. A slow improvement in completeness of registration with passing time was noted, although some states were known to have retrogressed after their admission to the registration area.

These indirect tests of underregistration advertised the need for a major direct test. A nationwide birth registration test was made in conjunction with the 1940 United States census. The census enumeration was used to provide a representative sample of births with which to search the registration files. A rather full exposition of the procedures followed is presented because it is believed they were carefully planned and carried out. Thus, they can be a guide to others who wish to undertake similar research. The following description is quoted from Grove (see basic readings).

"During the census enumeration in 1940, the enumerators were instructed to fill out a special infant card for all infants who were stated to have been born during the period December 1, 1939 to March 31, 1940, inclusive, and who were alive on April 1, 1940. At about the same time, all state and certain city bureaus of vital statistics prepared special copies of death certificates filed for children born during the period but dying before April 1, 1940.

"The basic problem then was to match each infant card and each copy of a death certificate with a copy of a birth certificate. Each unmatched infant card or death transcript would then be assumed to represent an unregistered birth.

"In order to match the records, the infant and death cards were first completely alphabetized by the last, first, and middle name of the child, by state of birth. The birth cards were alphabetized in the same way. These two groups of records for each state were then matched so far as possible. Infant and death cards which were unmatched after this operation, and the birth cards, also still unmatched, were then realphabetized by maiden name of mother, and the matching operation was repeated. A third alphabetization and matching operation was done by name of father. Finally, the remaining unmatched cards were sorted by county of birth and matched.

"Failure to match records which should have been matched in these operations was due, in most cases, to illegibility of the records, incorrect information on the records, or omission of pertinent information. In order to determine whether the remaining unmatched records could be matched if correct and complete information were given, three additional procedures were undertaken.

"First, a form letter was mailed to the parents of each child named on the unmatched infant or death cards. This letter requested the parents to supply all of the information needed to correct or complete

the original cards. If no reply was received from the parents within a reasonable time, and if the letter was not returned unclaimed, a second letter was sent to the hospital or institution in which the birth occurred, if one was indicated. Replies were received for approximately 74 percent of the letters sent. The number of unmatched infant and death records was thus further reduced by matching or by showing that the records should be discarded from the test because the births did not occur in the 4-month test period.

"The next step in the matching process was to check the remaining unmatched infant and death cards in the original birth certificate files of all state and certain city bureaus of vital statistics. There were several reasons for expecting that some records not previously matched could be matched in the files of original certificates...

"It was originally planned to have a complete field investigation in each state of all records unmatched after the check in the State Office files. Approximately one-third of the states made a complete or nearly complete check by correspondence or personal interview in time for inclusion of the results. Another one-third of the states completed a substantial part, but not all, of their field investigation. Finally, about one-third of the states did practically no field work.

"Time Standards--In order to establish a uniform test period throughout the country it was decided that birth certificates filed in the State Offices after July 31, 1940 would not be considered registered for purposes of the test. This rule had the effect of allowing 7 months for the filing of certificates for births occurring in December, 1939, 6 months for births occurring in January, 5 months for births in February and 4 months for births in March.

"The deadline was necessitated primarily by the stimulus to birth registration provided by the test itself. The letters sent to parents stimulated many to have their children's births registered. Also, the publicity given to the test by Federal, state, and local agencies tended to encourage registration... Since the maximum time allowed for filing certificates which are included in the annual tabulations... is about 16 months and the minimum is 5 months, the deadline used in the test tended to give a downward bias to the results for states in which the volume of late registration is substantial.

"Matching Standards.-- In order to assure maximum uniformity in the results for different states it was necessary to formulate and observe certain objective standards governing the matching of records. It was also important to set the standards at a point which would result in the closest approximation of the number matched to the true number which should have been matched. If the standards were set high enough to satisfy legal rules of evidence, very few incorrect matches would be made, but many true matches would be missed. Conversely, if

the standards were set too low many records would be matched which did not represent the same children. In this test an effort was made to set the standards between these two extremes so that the number of incorrect matches would approximate the number of true matches which were missed....The items considered in deciding whether two records matched were: name of child, name of father, maiden name of mother, date of birth, place of birth, and sex of child. If any four of these items, two of which were names, were essentially identical, the records were considered to refer to the same child and were matched, subject to verification. Cases which did not meet this standard but were believed to be true matches were referred to a supervisor for final decision."<sup>6</sup>

Representativeness of sample. Despite the fact that it was a complete coverage of information available from census and vital statistics records, the 1940 test was a sample of registered births, and the results are subject to errors of random sampling. In addition there was the possibility of bias arising from the sample of infant cards itself. There is a possibility that a disproportionate share of infants not enumerated by the census also are not registered at birth. The percentage completeness of birth registration has been computed on the assumption that the event of census nonenumeration and the event of birth nonregistration are independent or uncorrelated because no satisfactory information exists as to the degree of correlation that may be present.

Table 18-32. COMPLETENESS OF BIRTH  
REGISTRATION; UNITED STATES:  
DECEMBER 1 TO MARCH 31, 1940

Area and race	Percent of records matched
United States.....	92.5
White.....	94.0
All other.....	82.0
Cities of 10,000 or more ...	96.9
White.....	97.4
All other.....	91.5
Cities under 10,000 and rural	88.0
White.....	90.2
All other.....	76.6

Source: Grove, Robert D., "Completeness of Birth Registration in the United States"; United States Bureau of the Census, Special Reports, Vol. 17, No. 18, 1943.

Table 18-32 gives the percentage completeness of birth registration for the United States, by race and by population size groups for 1940. The proportion of births found to be registered are termed "Completeness of registration factors." Registered births for the year 1940 generally are corrected for underregistration by dividing them by the factors. In two reports entitled Studies in Completeness of Birth Registration (one for the nation and states, and one for counties and cities of 10,000 or more), underregistration factors were published for each state, county, and city of 10,000 or more. These factors refer to births by place of occurrence. A special retabulation was later made to make these same factors available for 1940 by place of residence.

In the period following 1940 there was a continued improvement in birth registration. Several factors were undoubtedly responsible.

(1) The 1940 birth registration test gave wide publicity to the deficiencies of the registration system. Places where registration was poor were identified by name. This was undoubtedly a stimulus for many states and counties to improve their systems, and it provided a guide for locating the spots where improvement was most needed.

(2) During the 1940's many firms producing materials for defense and war refused to hire persons who could not supply proof of citizenship. A birth certificate was the best evidence of citizenship. This emphasized among the public the importance of being able to certify the facts of birth and probably caused a higher proportion to make certain that their children were properly registered at birth.

(3) A disproportionately large share of unregistered births occur outside of hospitals. In the 1940 test it was found that 98.4 percent of the births that occurred in hospitals were registered, while only 85.5 percent of births occurring outside hospitals were registered. Between 1940 and 1944 the proportion of births occurring in hospitals increased from 55.8 percent to 75.6 percent. Even if hospitals continued to register births that occurred there at the same rate as 1940, and if births occurring outside hospitals were registered no more completely than in 1940, the completeness of birth registration would have improved. This fact was pointed out by Moriyama, who used the changing proportion of births occurring in hospitals to compute underregistration factors for single years between 1935 and 1940 (as well as temporary estimates for 1941-44). Because these estimates are extremely useful in dealing with birth statistics for the earlier years they are reproduced below:

<u>Year</u>	<u>All races</u>	<u>White</u>	<u>Nonwhite</u>
1940	92.3	94.0	81.3
1939	91.9	93.6	81.1
1938	91.6	93.4	80.8

<u>Year</u>	<u>All races</u>	<u>White</u>	<u>Nonwhite</u>
1937	91.3	93.1	80.4
1936	91.1	92.8	80.2
1935	90.5	92.5	78.7

In his article (cited in the basic readings) Moriyama gives information that can be combined with published statistics for these earlier years to compute underregistration factors for individual states.

A second nationwide test of birth registration completeness was made in conjunction with the 1950 census. The same general approach as that used in the 1940 census was adopted. However, the study was broadened to include the study of infant underenumeration (see below). As a part of the 1950 census enumeration, an infant card was filled out by the enumerator for each infant born between January 1 and March 31, 1950. Although this was a shorter time period than in the 1940 test, the birth rate had risen and the number of cases was larger in the new test than in the first. A unique aspect of the 1950 test was that much of the matching of birth certificates with infant cards was done by machine, using punched cards operated by highly skilled operators. When all possible matches had been made, a mail survey was made requesting parents (in some cases hospitals, welfare agencies, and other agencies who might have special records) to verify, correct, or complete information on the infant cards. Follow-up letters were sent in cases of nonresponse. This removed duplicates, records for children born outside the test period, and helped match additional records. Cards that remained unmatched after the mail survey were sent to registration offices in each of the states for searches against the individual birth record files. The state registrars and their staffs exhausted all possibilities for locating matching records. This included special investigations into other records involving adoption, confusion concerning name in situations involving illegitimacy, change in name, and enumeration of a child in the home of a grandparent or other relative. All additional matches made by the state offices were reviewed by the National Office of Vital Statistics to ensure comparability of test results.

Final measures of birth registration completeness were computed by dividing the total number of infant cards matched with birth records by the total number of infant cards, matched and unmatched.

In the 1940 enumeration test, a sample of records for infant deaths had been included. Death records were excluded from the 1950 project to simplify the operation. This has virtually no effect on the measures of completeness in view of the few death records involved. The following quotations from the reference to Shapiro and Schachter (see basic readings) summarize the results of the test.

"The 1950 birth registration test indicated that 97.9 percent of the

infants born in the early part of that year had birth certificates on file in vital offices. In 24 states and the District of Columbia, birth registration completeness was 99.0 percent or more, and in only 7 states was it lower than 95 percent.

"Seven out of eight infants included in the test were born in hospitals, and all but a few of the hospital births were registered. For births delivered at home, however, registration was not nearly as complete. Nationally, only 88.2 percent of these births were registered, and in some states the proportion was considerably lower.

"About two-fifths of the births in the test occurring at home were attended by midwives, relatives, or neighbors. These attendants registered 84.8 percent of the births they delivered compared with 90.7 percent for physicians attending home deliveries. Nonphysicians were used far more often in the South Atlantic, South Central, and Mountain Divisions than in other parts of the country. In a number of states these attendants took care of more births at home than did physicians, and in some areas they had a better record of registration.

"Of the white births in the test, 98.6 percent were registered, compared with 93.5 percent of the nonwhite. A closer examination of the situation indicates that there was no difference between the two race groups in registration completeness of births "at home" and only a slight difference with respect to the "in hospitals" births. However, when hospital and nonhospital births are combined, registration is found to be more complete in the white group than in the nonwhite because of the more frequent occurrence of white births in hospitals.

"A comparison of results from the 1940 and 1950 registration tests shows that substantial gains were made during the intervening years. About four-fifths of the increase is explained by the trend toward use of hospital facilities for obstetrical care. The remaining portion of the improvement was due to moderate increases in registration of births that occurred in hospitals as well as at home. Registration improved in virtually every state. Large gains were made in most of the Southern states, nearly all of which were well below the 90 percent point in 1940. Registration in the nonwhite races improved considerably during the 1940-50 decade. As a result, the wide difference in registration completeness between the white and the nonwhite groups that existed in 1940 was substantially reduced. Nonwhite registration improved by 2.0 percent for hospital births and by 14.2 percent for births delivered at homes. The importance of the latter change is indicated by the fact that even in 1950, nearly half of the nonwhite births occurred at home. Promotional efforts of many of the Southern states were directed primarily toward this group."<sup>7</sup>

The results of the 1950 registration test are summarized in two publications entitled, Birth Registration Completeness in the United

States and Geographic Areas, 1950: Part I--Data for each State and Part II--Data for each Country and Urban Place of 10,000 or More. Table 18-33 summarizes some of the major completeness of registration factors. In addition to showing the completeness of registration factors for many areas, the test results of 1940 also provided this information by age of mother, nativity, order of birth, and other characteristics.

Use of the Completeness of Registration Factors in Demographic Research. With the results of the enumeration tests of 1940 and 1950 it is possible to correct for underregistration the registered births for any area for any year, from 1940 to the present. The procedure consists of establishing by interpolation a completeness of registration factor for the place and year desired, and then to divide this factor into the registered number of births.

Table 18-33. PERCENT REGISTRATION COMPLETENESS FOR BIRTHS IN HOSPITALS AND NOT IN HOSPITALS, BY RACE AND URBAN-RURAL RESIDENCE; UNITED STATES: JANUARY 1 TO MARCH 31, 1950

Race and area	Total	In hospital	Not in hospital
All races.....	97.9	99.4	88.2
Urban.....	98.9	99.5	89.8
Rural.....	96.5	99.2	87.7
White.....	98.6	99.5	86.2
Urban.....	99.3	99.6	89.5
Rural.....	97.6	99.3	87.9
Nonwhite .....	93.5	98.2	88.2
Urban.....	96.4	98.6	90.1
Rural.....	90.1	97.0	87.4

Source: National Office of Vital Statistics. "Birth Registration Completeness in the States and Geographic Areas, 1950 Part I, Data for each State," Vital Statistics - Special Reports, Selected Studies, Sept. 21, 1954, table D.

The most accurate interpolation for years between 1940 and 1950 is obtained by making a linear interpolation separately for the completeness of enumeration of hospital and nonhospital births by color (white and nonwhite). This provides four interpolated values for the area. These factors are then applied to the statistics for births subdivided into the corresponding four groups. However, statistics are published by race for hospital and nonhospital births only for states and large places. For other areas it is necessary to make a linear interpolation between the 1940 and the 1950 underregistration factor for total births

(making a separation for white and nonwhite if possible), without reference to occurrence or nonoccurrence in a hospital.

Since 1950 the completeness of registration has improved, but primarily because of the greater proportion of infants born in hospitals. By using the 1950 factors separately for hospital and nonhospital births, results that are very nearly correct can be obtained. By 1960 the proportion of births occurring in hospitals had reached such a high level that correction for underregistration has almost no effect upon fertility rates.

The completeness of registration factors for 1950 are computed and reported in two ways by place of occurrence of the birth and by place of residence. The demographer should be careful to apply the factors that pertain to the type of vital statistics he is using. Correction factors for place of occurrence should not be applied to statistics that are tabulated by place of residence.

18-18b. Underenumeration of Children. The census enumeration tends to undercount children under ten years of age. The most serious underenumeration is for infants under one year of age. This deficiency has been noted as a characteristic of many censuses in different parts of the world. It may be observed directly in a variety of ways. In some cases, the population enumerated as under 5 years of age in one census has been less than the number enumerated as 10 to 14 years old in a census taken a decade later, with allowance for deaths and migration in the interval. The population enumerated as under one year old has sometimes been less than the number enumerated at ages 1, 2, 3, or 4 years, even though there has been no decrease in the number of births. The child population generally has tended to be less than the number expected from births (corrected for under-registration, deaths, and net migration).

Because the statistics for these younger ages (especially 0-4 years inclusive) are frequently used to measure fertility among a population, this error is of considerable importance for fertility analysis. Before the techniques described in Sections 18-10, 18-11, 18-12, and 18-13 can be employed, it will be necessary in many cases to correct the statistics for children under 5 for underenumeration. The information available about this error in the United States Census, and methods for correcting data to compensate for it, are the subject of this section.

Richard A. Hornseth of the United States Bureau of the Census has made careful estimates of the undercount of children and the misclassification of people of all ages in both the 1940 and 1950 Censuses. The results are published in the Current Population Reports cited in the basic readings. His data for the age group 0-4 years are summarized in Table 18-34. These estimates were obtained by (a) estimating the births that occurred between April 1, 1945, and April 1, 1950 (cor-

rected for underregistration); (b) subtracting deaths that occurred in this group of births before April 1, 1950 (which involved transforming the data for deaths to period of birth, correcting deaths to infants under one year for underregistration); and (c) making an allowance for net immigration from abroad. These figures indicate the extent to which it is impossible to reconcile the 1950 and the 1940 Census counts with the registration data for infants and immigration records. They estimate what proportion of the true number of children 0-4 were actually enumerated by the censuses. Such proportions are termed "completeness of enumeration factors." Since the completeness of registration factors employed for correcting birth and death data were derived from samples, and are therefore subject to sampling error as well as possible bias, it is entirely possible that not all of this "residual" or discrepancy is due to census underenumeration. By the same token, the statistics for immigration may be subject to some error. Only by assuming that the errors in the birth, death, and migration data are all zero after correction, or that they cancel each other, is it possible to assume safely that the percentages of Table 18-34 are true measures of underenumeration. This caution applies with special emphasis to the factors for nonwhites. Since the figures are a residual, all errors made in estimating the other components are contained in those numbers, and a comparatively small difference in the underregistration factors could effect a large change in the estimated amount of underenumeration. Since these estimates represent a most careful review of the available evidence, they are the best information at hand concerning the extent of underenumeration of children under 5 years of age.

Table 18-34. HORNSETH ESTIMATES OF CENSUS UNDER-ENUMERATION AND COMPLETENESS OF ENUMERATION FACTORS FOR CHILDREN 0-4 YEARS OF AGE: 1940 AND 1950

Sex and color	1940		1950	
	Estimated number not enumerated	Percent completeness of enumeration(a)	Estimated number not enumerated	Percent completeness of enumeration(a)
	(1)	(2)	(3)	(4)
Total.....	362,528	90.8	802,446	94.2
Male.....	456,669	90.5	436,932	93.8
Female.....	405,859	91.2	365,514	94.6
White.....	626,377	92.2	590,868	95.1
Male.....	332,224	91.9	326,695	94.7
Female.....	294,153	92.6	264,173	95.5
Nonwhite.....	236,151	81.9	211,578	88.1
Male.....	124,445	81.1	110,237	87.7
Female.....	111,706	82.8	101,341	88.5

(a) Obtained by adding estimated number not enumerated to the census count and computing percent census count of adjusted total.

These data show that the rate of underenumeration in 1940 was higher than in 1950. In both years, the rate for male infants was higher than that for female infants.

For the Infant Enumeration Study, 1950, Grabill computed completeness of enumeration factors for the first five years of life as follows:

0-1 years	89.0
1-3 years	95.2
3-5 years	98.8
Total	95.3

Thus, the underenumeration seems to be most serious in the first year of life and to approach zero at ages 3 and 4.

For the 1960 Census of the United States a special effort was made to improve age statement and coverage for young children, with the result that infants were enumerated with almost the same precision as adults.

For the report, Differential Fertility, 1940 and 1910, the United States Bureau of the Census computed completeness of enumeration factors for each state, by urban-rural residence by color for 1940, and for each region, by urban-rural residence by color for 1910. These factors are reported in Tables A-1 and A-2 of that publication. In making use of them, it must be remembered that they state what proportion of the true population has been enumerated. Therefore, they must be divided into the census counts to obtain the correct results. As yet, the Bureau of the Census has not released completeness of enumeration factors for each state comparable to those computed for 1940. One of the outstanding characteristics of the 1940 tabulation was that the variation among the states in degree of underenumeration was comparatively small. Because of this fact, a serious error would probably not result from using the factors of Table 18-34 to adjust state and local census totals, by sex and color. In 1940 it was shown that underenumeration was much more severe in rural farm areas than in urban areas. It seems plausible that a similar situation existed at the time of the 1950 Census, although perhaps to a smaller extent. The factors from the infant enumeration study could be linked, by ratios, to provide completeness of enumeration factors for the 0-4 age group for rural and urban areas in each state, by color.

The Infant Enumeration Study of 1950. In order to gain more information about the problem of underenumeration of infants, the United States Bureau of the Census and the National Office of Vital Statistics cooperated, at the 1950 Census, to conduct an infant enumeration test. This test was designed along lines similar to the infant birth registration test. After the matching operation described for the birth

registration test had been carried out, the Bureau of the Census took a random sample of the birth records for which a census infant card had not been prepared (excluding illegitimate births and certain other classes of births) and made an intensive search of the census schedules to determine whether or not those infants actually had been missed in the enumeration. If a search of the census schedules failed to indicate that the infant had been enumerated (sometimes the enumerator recorded infants on the official schedules, but failed to fill out a special infant card), a letter was sent to the parents of the child asking for the address of the child at the time of enumeration, with a request for the possible reasons for the missing of infants in the census. Follow-up letters were sent in case a reply was not received. With the additional information supplied by the parents, a renewed check of the census files was made in an effort to locate the infant in the census schedules. On the basis of this search, the infants were classified either as definitely enumerated, definitely missed, or in one of three (proportionately small) categories: "probably enumerated," "probably missed," and "enumeration status uncertain." By an estimating procedure, an "adjusted estimate of percent enumerated" was obtained for many different classes of the population and for many different areas. In the census publication, Infant Enumeration Study, 1950, statistics of enumeration of infants 0-3 months old are shown for each state by urban and rural residence, and by regions, for sex, order of birth, month of birth, age of mother, years of school completed by mother, major occupation group of father (all of the above cross-classified by color and urban-rural residence). Table 18-35 summarizes some of the basic completeness of enumeration factors.

Table 18-35. ADJUSTED ESTIMATE OF PERCENT ENUMERATED FOR INFANTS BORN IN FIRST THREE MONTHS OF 1950, BY COLOR AND AREA

Area and color	All infants	Male infants	Female infants
Total	96.4	96.4	96.5
White.....	97.1	97.1	97.1
Nonwhite.....	91.4	91.3	91.5
Urban	96.8	96.8	96.9
White.....	97.5	97.4	97.6
Nonwhite.....	91.8	91.6	92.0
Rural non-farm	96.7	96.8	96.6
White.....	97.1	97.2	97.0
Nonwhite.....	92.2	92.3	92.1
Rural farm	94.7	94.7	94.7
White.....	95.8	95.9	95.8
Nonwhite.....	89.9	89.7	90.2

Source: U.S. Bureau of the Census, Infant Enumeration Study, 1950, table 3.

The surprising result of this test is that 96.4 percent of the infants were enumerated. Yet it was demonstrated above that the census count of infants under one year was 11.0 percent too small. These two statements can be reconciled only by assuming that the apparent drastic undercount of infants under one year of age actually consists of a tendency to misreport age. It may take the form of reporting infants that are 10 or 11 months old as being one year old.

The mailed questionnaire brought explanations for underenumeration from the parents. In about 82 percent of the cases in which infants were classified as definitely or probably missed in the census, the parents were also missed. In the 18 percent of cases where the parents were enumerated but the infant was missed, the reasons given for the oversight included: (a) a neighbor gave incomplete information; (b) the family did not think infants were to be reported; and (c) infant died between April 1 and the time of the census enumeration. In only about 5 percent of the cases did the family forget to report the infant or think infants were not to be reported. This discredits the belief, which demographers previously held, that mothers have a tendency to forget to report their young babies to the enumerator.

Additional information about underenumeration of children under 5 years of age is provided by the Post-Enumeration Survey conducted as a quality check by the United States Bureau of the Census. This survey revealed that there is a considerable tendency (probably unintentional) to misreport the age of children and that the undercount for ages under 5 is a net difference between a rather common tendency to report children as being older than they actually are. A smaller reverse tendency to report some children as being younger than they actually are and a moderate tendency for children to be missed (either as individuals or as members of families which are missed) were also noted.

Thus, the underenumeration of infants under three months of age is seen to be only a part of the general problem of underenumeration of children under 5 years, and not a special problem in itself. The special problem seems to consist of the misreporting of children's ages, especially at the level of about one year.

18-18c. Underreporting of Children Ever Born. Children-ever-born data tend to be of low precision in many countries, although for recent censuses in the United States they are of good quality.

If children-ever-born data are to be used only to measure differentials in fertility, it is not necessary to correct the data for underreporting; it may be presumed that underreporting is equally serious among all subgroups. In general, this will be a conservative procedure because it will tend to understate the size of fertility differentials.

If children-ever-born data are to be used to measure fertility levels, however, they must be adjusted, for they are almost certain to under-

state cumulative fertility by a wide margin. Yet it is not easy to estimate the amount of this understatement. P. K. Whelpton explored this problem more than any other demographer, for he used CEB data to build up his parity rates for real cohorts. His methodology for estimating the margin of error from the censuses of 1940 and 1950 is a complex one. The reader is referred to his monograph Cohort Fertility for the details.

## FOOTNOTES

<sup>1</sup>Donald J. Bogue and James Palmore, "Some Empirical and Analytic Relations Among Demographic Fertility Measures, with Regression Models for Fertility Estimation," Demography, Vol. 1 (1964), pp. 316-38. In these equations and in Tables 18-3 and 18-4, GFR is expressed as births per 1,000 women aged 15-49 years.

<sup>2</sup>Donald J. Bogue. Principles of Demography. (New York: John Wiley and Sons, Inc., 1969).

<sup>3</sup>El Badry, M.A., "Effect of Errors in Reporting Childlessness Upon Measuring Fertility," Journal of the American Statistical Association.

<sup>4</sup>Wilson H. Grabill and Lee-Jay Cho, "Methodology for the Measurement of Current Fertility from Population Data on Young Children," Demography, Vol. 2 (1965). An even more complete and refined statement of these procedures will be found in Chapter 9 ("Methodology") of Lee-Jay Cho, Wilson H. Grabill and Donald J. Bogue, Differential Current Fertility in the United States. This chapter is the most detailed exposition of the "own children" methodology available.

<sup>5</sup>Op cit., Chapter 9, pp. 319-21. Procedures for making adjustments for individual states or other areas smaller than the entire nation are also described in this chapter.

<sup>6</sup>Grove, Robert D. "Studies in the Completeness of Birth Registration," December 1, 1939 to March 31, 1940. U.S. Bureau of the Census. Vital Statistics, Special Reports, Vol. 17, Vol. 18, 1943.

<sup>7</sup>Sam Shapiro and Joseph Schachter, "Birth Registration Completeness in the United States and Geographic Areas, 1950," Vital Statistics, Special Reports, Selected Studies, September 21, 1964.



